(12)

Technical Report NAVTRAEQUIPCEN 81-C-0105-3

20000803019

AD-A145 218

CONTENT, VARIETY, AND AUGMENTATION OF SIMULATED VISUAL SCENES FOR TEACHING AIR-TO-GROUND ATTACK

G. Lintern, K. E. Thomley,
B. E. Nelson, and S. N. Roscoe

Canyon Research Group, Inc.
741 Lakefield Road, Suite B
Westlake Village, California 91361

July 1984

Interim Final Report for Period
1 September 1981 - 31 August 1984

NAVAL TRAINING EQUIPMENT CENTER
ORLANDO, FLORIDA 32813

84 08 30 023

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NAVTRAEQUIPCEN 81-C-0105-3	2. GOVT ACCESSION NO.: AD-A145 28
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): CONTENT, VARIETY, AND AUGMENTATION OF SIMULATED VISUAL SCENES FOR TEACHING AIR-TO-GROUND ATTACK	5. TYPE OF REPORT & PERIOD COVERED: Interim Final Report 1 Sep 81 - 31 Aug 84
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): G. Lintern, K. E. Thomley, B. E. Nelson, and S. N. Roscoe	8. CONTRACT OR GRANT NUMBER(s): N61339-81-C-0105
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Canyon Research Group, Inc. 741 Lakefield Road, Suite B Westlake Village, California 91361	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 0785-2P3
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Training Equipment Center Orlando, Florida 32813	12. REPORT DATE: July 1984
	13. NUMBER OF PAGES: 74
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Aircraft Simulation	Skill Acquisition	Flight Training
Human Factors	Computer Generated Displays	Augmented Feedback
Transfer of Training	Training Displays	Scene Content
Instructional Techniques	Visual Simulation	Variety in Training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Visual Technology Research Simulator was used for quasi-transfer-of-training study in which 32 military pilots were taught to deliver bombs from a 30-degree dive. Sixteen of the pilots had a moderate amount of prior bombing experience (approximately 60 bombing runs) and the remainder had none. The pilots were given 80 training trials in the simulator under specific training conditions. Three factors were manipulated in training; those being level of detail in the visual scene, number of visual scenes, and augmented feedback in the form of artificial visual guidance. Differential transfer

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

effects were assessed on the basis of performances on 30 transfer trials in the simulator. The transfer phase used a variety of visual scenes and varying levels of detail but no augmented feedback. All subjects flew the same set of conditions in the transfer phase.

Scene content had an unexpected, but strong and consistent effect on performance and on differential transfer. A Landscape scene that contained buildings, roads, and rectangular fields was generally better than a schematic Grid Pattern for both training and transfer. The results of this experiment did not clearly isolate specific scene features that contributed to this effect. However, some likely candidates were identified and these will be investigated more closely in future experiments.

The scene-content issue is one of the most crucial for modern training simulators. These data are the first to show that scene content affects learning of flight skills. Further research to identify visual features that do impact learning is essential. In the meantime, simulator training of air-to-ground attack should be conducted with visual scenes that at least have features similar to those of our Landscape.

Variety was raised as a training issue, specifically because modern simulators can provide enormous variety at little additional cost. Scene variety in training did not generally benefit transfer, and there is a distinct possibility that it can interfere with early learning. However, transient disruptions in performance at transfer suggested that brief experiences with a wider range of scenes towards the end of a constant training regimen would be useful.

Augmented feedback proved to be a potent instructional variable, but one that showed complex effects. It helped inexperienced pilots with their dive pitch control, and helped both the inexperienced and more experienced pilots with their dive altitude control. The data further indicated that augmented feedback helped the more experienced pilots with their longitudinal bomb miss distance. Thus, the effects of augmented feedback are pervasive and progressive. It ~~would~~ appear to be useful at least for primary and intermediate instruction.

There were several interactions of pilot experience with the experimental variables. In general, the inexperienced pilots suffered most from limited scene content, and gained most from augmented feedback. Nevertheless, the moderately experienced pilots were also affected by these variables. Thus, there is no evidence in this report that pilots with no experience in air-to-ground attack should be treated differently during training to pilots who have some experience in air-to-ground attack.

SUMMARY

Thirty-two military pilots were taught manual (not computer aided) bomb deliveries in the Visual Technology Research Simulator (VTRS). Scene content, scene variety, and augmented feedback were manipulated between groups during training. All pilots were subsequently tested in the VTRS on an identical series of bombing trials.

Scene content had an unexpected but strong and consistent effect during training and transfer. A Landscape scene that contained buildings, roads and rectangular fields was generally better than a schematic Grid Pattern for both training and transfer, although the Grid Pattern proved to be superior on some performance measures. While it was not generally possible to determine scene features that contributed to this effect, some likely candidates were identified.

The most intriguing observation on scene content was that pilots who learned the task with the Landscape could later perform well with the Grid Pattern. However, those trained on the Grid Pattern never exhibited the high level of performance shown by those trained on the Landscape. Some of the features of the Landscape scene seem essential for early learning. Their value for later training could not be determined.

Variety was raised as a training issue specifically because modern simulators can provide enormous variety at little additional cost. Scene variety in training did not generally benefit transfer, and there is a distinct possibility that it can interfere with early learning. However, transient disruptions in performance at transfer suggested that brief experiences with a wider range of scene towards the end of a constant training regimen could be useful.

Thus, we recommend that air-to-ground attack be taught initially in a simulator with only one scene and run-heading, but that a variety of scenes and run-in headings be introduced just prior to transfer. A similar approach might be used in the aircraft if the option exists. In addition, it may be possible to simulate attacks on actual targets. This might increase the effectiveness of a pilot's first pass at that target. The effect on training of varying other factors of the task could also be examined. Variations in environmental conditions and ordnance delivery modes are good candidates for further research.

Augmented feedback proved to be a potent instructional variable, but one that showed complex effects. It helped inexperienced pilots with their dive pitch control, and helped both the inexperienced and more experienced pilots with their dive altitude control. The data further indicated that augmented feedback helped the more experienced pilots with their longitudinal bomb miss distance. Thus, the effects of augmented feedback are pervasive and progressive. It would appear to be useful at least for primary and intermediate instruction.

There were several interactions of pilot experience with the experimental variables. In general, the inexperienced pilots suffered most from limited scene content, and gained most from augmented feedback. Nevertheless, the more experienced pilots were also affected by these variables. Thus, there is no generalizable evidence that pilots with no experience in air-to-ground attack should be treated differently during training from pilots who have some experience in air-to-ground attack.

The experimental design issue of statistical power was also considered. Statistical power remains a problem area. Despite our best efforts, power remained low on some important performance measures. The most obvious means of increasing power, that being to increase the number of experimental subjects, will often be considered too expensive. The options of finding satisfactory covariates and of improving performance measures should be pursued.

PREFACE

Dr. Stanley C. Collyer and Dr. Dennis C. Wightman were the contracting officer's technical representatives for this research, and assisted with planning and interpretation. Mr. Walter S. Chambers, facility manager at the VTRS, and many other members of the support staff assisted with conduct of the experiment. Daniel J. Sheppard and Ronald S. Mauk of the Essex Corporation assisted with data collection and analysis. Dr. Jack A. Adams, University of Illinois; Dr. William P. Dunlap, Tulane University; Dr. Marshall B. Jones, Pennsylvania State University; Dr. Walter Schneider, University of Illinois; and Dr. Kent A. Stevens, University of Oregon, advised during planning and interpretation.

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS

(cont'd)

LIST OF ILLUSTRATIONS

LIST OF TABLES

(cont'd)

LIST OF TABLES (cont'd)

SECTION I

INTRODUCTION

Air-to-ground weapon delivery is currently taught on several ranges throughout the U.S. and with dummy targets at sea. None of the available ranges provides realistic representations of wartime target areas, and all have only a small number of targets that each pilot must attack repeatedly in training. A few ranges with more numerous and realistic targets exist but these tend to be used for advanced training. In basic weapon delivery training especially, practice is generally restricted to one target in one training area.

PROBLEM

Simulation affords the only feasible opportunity to provide students with a wide range and well-ordered schedule of targets and environments. However, there is very little information that can guide us in selecting the appropriate number or type of visual scenes for teaching air-to-ground attack. It is unknown whether several visual scenes would provide better training than only one. It is also unknown whether scenes that are provided should be realistic and rich in scene content, or simple and schematic. Resolution of these issues would impact both cost and training effectiveness of training simulators.

If differences between high and low realism were discovered, it would be useful to identify the visual features that led to those differences. Other issues that are pertinent to the design of visual scenes for teaching air-to-ground attack are whether learning can be enhanced with augmenting visual guidance and whether pilot experience modifies the effects of any of these variables. In the investigation reported here, we explored the effects of these variables on transfer.

RESEARCH ISSUES

SCENE CONTENT. The previous research on realism or scene content suggests that schematic scenes have substantial training value. Lintern (1980) has shown that some landing skills can be acquired with a visual representation that is limited to a runway outline, a horizon, and a centerline. Westra (1982) has compared training with a solid surface representation versus training with a point-light outline in carrier-landing research, and has shown that scene content has only a small and transitory effect on differential transfer. Thus, students can apparently learn to land on carriers with a light-point representation of the carrier and the additional benefit of solid surfaces appears to be marginal. However, flight tasks vary considerably in

their information requirements, so that experimental results obtained with one task should not be generalized automatically to other flight tasks. Thus, it is essential to reexamine the scene-content issue in the air-to-ground bombing context.

SCENE VARIETY. Basic research data generally support the notion that varied training benefits transfer. Although training performance is generally poorer, transfer to a different version of the same task has been shown to be better after varied training with substitution tasks (Dashiell, 1924), memory tasks (Duncan, 1958), problem solving tasks (Morrisett and Hovland, 1959) and target identification tasks (Schneider and Fisk, 1982). However, this advantage may not appear if there is no common element between training and transfer tasks (Crafts, 1927) or if very little training is given on each of the tasks that make up the varied set (Adams, 1954).

An air-to-ground bombing task could be varied either by changing the information that is available in the visual scene, or by using different environmental conditions to force changes in motor responses. Other variations such as modification of the bombing pattern, or delivery of different ordnance, would vary both visual information and motor responses. The theoretical development by Schmidt (1975) seems to suggest that any type of variation is pertinent to this issue. He has postulated that skill learning involves development of generalized schema and that varied training will strengthen the appropriate schemata.

In contrast, Ellis (1969) has emphasized the visual aspects of skill learning. He has argued that varied training produces good attentional habits by forcing students to attend closely to the available stimuli. He suggests that these good habits transfer. In the same discussion Ellis (1969) also proposes a related but independent hypothesis that varied training develops discrimination of relevant and irrelevant cues. This latter notion is consistent with a well-supported argument by Gibson (1969) that the acquired ability to discriminate relevant and irrelevant cues is important to perceptual learning.

A variation of scene content was chosen primarily because of the emphasis in this program on visual display issues. This choice seemed appropriate in the light of extensive data from Schneider and his associates that are relevant to information processing manipulations. While Shiffrin and Schneider (1977) emphasized the value of consistency in stimulus presentation for learning target identification, Schneider and Fisk (1982) have subsequently shown that variation in the training task is advantageous if subjects are transferred to a similar but nonidentical task. On the other hand, a comprehensive review by Shapiro and Schmidt (1982) has failed to show any consistent training advantage for variations in motor output, especially for adult subjects.

The laboratory findings from Schneider and Fisk (1982) suggest that variation across the practical range of a task, especially with regard to dimensions of perceptual processing, will produce better transfer to a task in which these variations occur naturally. In the air-to-ground context it would be consistent with the notion of learned discrimination of relevant and irrelevant cues (Ellis, 1969; Gibson, 1969) that students who practice with more than one visual scene might ignore cues that are useful but specific to only one of the visual scenes used. Thus, varied training would encourage use of cues that are common to different visual scenes so that transfer to any new scene would be more likely to occur without serious disruption.

AUGMENTED FEEDBACK. The use of augmented-feedback cues to aid acquisition of tracking skills has a long history in experimental psychology. This literature has been reviewed by Lintern (1978) and summarized by Lintern and Roscoe (1980). In brief, these data indicate that augmented feedback can aid performance during training and that the performance advantage will sometimes be retained when the supplementary cues are withdrawn. The data from approximately 30 experiments of this type are consistent with the hypothesis that the performance advantage will be retained only if the experimental subjects do not become dependent on the supplementary visual cues.

Lintern (1980) suggested that an off-target presentation, in which the cues appeared only when the subject exceeded specific error limits, would avoid dependency. He obtained data from simulated aircraft landings to support this hypothesis. Hughes, Paulsen, Brooks, and Jones (1979) tested an augmented feedback manipulation in air-to-ground bombing. Some of their subjects learned the task with the addition of a bomb-impact-predictor cue. There was no advantage shown for that type of training. However, their predictor cue did not facilitate performance during training, nor was there any special attempt to avoid dependency on it. Lintern (1978) argued that both of these considerations are important in the use of augmented feedback for acquisition of tracking skills.

Air-to-ground weapon delivery appears to be similar to the landing task at least in that pilots must follow a specific flight path for which the visual cues are poorly defined. These are the particular conditions that would seem to provide an opportunity for augmented feedback to be effective.

SUBJECT EFFECTS. Individual differences have been largely ignored in flight simulation research. Nevertheless recent work has shown that individual differences account for a substantial portion of the experimental variance, and usually outweigh the effects of the major experimental variables (Westra, Simon, Collyer & Chambers, 1982; Westra, 1982). This remains true even with experienced subjects who are selected from a

population that is generally regarded as homogeneous (Westra et al., 1982). An effort to account for individual differences and to ascertain whether they interact with other experimental variables seems worthwhile. In one successful application of this approach, high- and low-aptitude subjects have been shown to respond differently to part-task procedures for teaching carrier landings (Wightman, 1983).

Experience is one dimension of individual difference that is an important issue in military aviation because flight training must be conducted through the entire range from the undergraduate pilot to the seasoned aviator. Of particular concern for this study is that air-to-ground bombing is taught at novice, transition, and maintenance levels. Instructional techniques that are good for the novice may not be effective for transition training with a pilot who is going through his second or third course in weapon delivery, or with the experienced pilot who is maintaining well-learned skills. Experience was considered an important variable in this study for its possible modifying effects on the other experimental variables, although for practical reasons it was necessary to restrict experience levels to the low end of the range.

EXPERIMENTAL APPROACH

A within-simulator transfer-of-training study was conducted in which military pilots with little or no air-to-ground experience were taught a dive bombing task. Training was conducted with either one or two visual scenes, and pilots were subsequently tested on both visual scenes, plus one that had not been used in training. This procedure was used to ascertain 1) whether there was any loss in transfer to a new scene, and 2) whether the varied experience with two scenes permitted better transfer to a new visual scene. The same start point and run-in line were used throughout training, but a new start point and run-in line were added in transfer to check whether pilots referred to specific scene features while performing the task.

SCENE CONFIGURATIONS

The two training scenes differed almost to the maximum extent possible in the Visual Technology Research Simulator (VTRS). One showed topographic features such as roads, towns, outlying buildings, and colored fields (Figure 1, Landscape). The alternate scene consisted of a white grid pattern on a green background with a white cube as a target (Figure 2, Grid Pattern). Of the subjects who were trained on only one scene, half flew the Landscape scene, and the remaining half flew the Grid Pattern. This enabled a test of the relative training value of the two scenes.

The additional transfer scene can best be described as a River Valley. Two mountain ranges, up to 4,000 feet high and

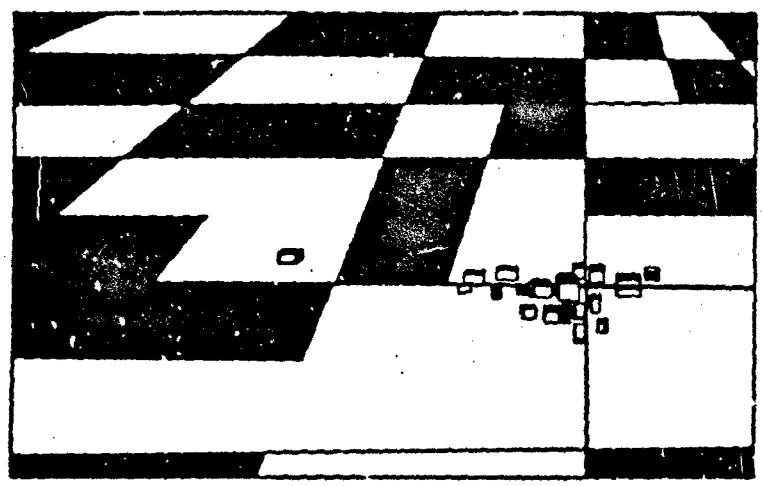


Figure 1. Black and White Representation of the Multi-colored High-detail Landscape (LS) Used in Training and Transfer. The isolated building slightly left of center of the figure was the designated target. A horizon and sky were also a part of this scene.

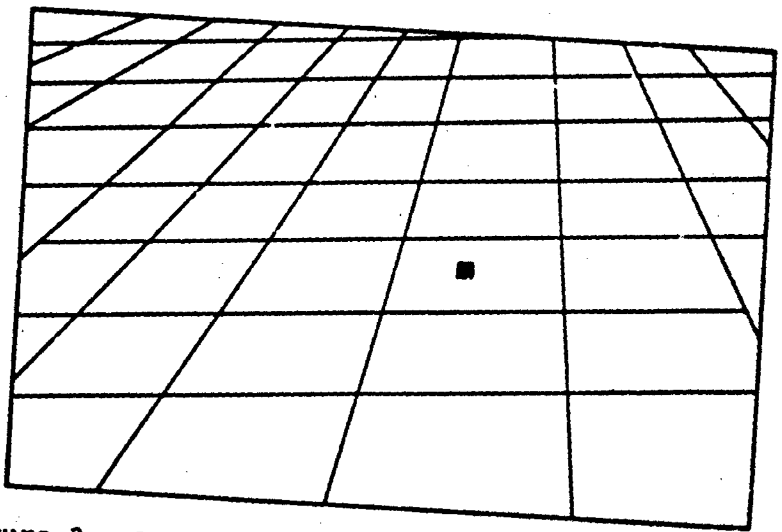


Figure 2. Low-detail Grid Pattern (GP) Used in Training and Transfer. The square to the right of center represents a white cube that was the designated target. The grid lines were white, and the background was green. A horizon and sky were also a part of this scene.

approximately a half-mile apart, paralleled a river with two bridges, a small town, and some small vessels. The white cube from the Grid Pattern was on the river surface as a target (Figure 3).

ADAPTIVE AUGMENTATION

The efficacy of supplementary visual cues was also tested. Supplemental visual cues were used to identify the flight path through the bombing pattern. These cues were presented in an adaptive mode; that is, they were visible only when prespecified error limits were broken. This ensured that, as the subject learned to make the correct response, the assistance provided by the augmentation would decrease. Thus, dependency on it would be avoided and subjects should be able to transfer to the nonaugmented condition without disruption of behavior patterns learned in the augmented condition.

TASK SELECTION

Task selection was the focus of a major preexperimental effort. The rationale for selecting a manual (not computer-aided) bomb delivery from a 30-degree dive was developed by Vreuls and Sullivan (1982), and was based on the four general criteria of 1) operational relevance, 2) consistency with current training practices, 3) suitability for examining visual simulation issues, and 4) feasibility within the current VTRS configuration. The constraints of the last criterion included T-2 aircraft dynamics, an A-4 weapon control panel, and simulation of an optical bombsight. Bombs, rockets, and guns were available. The 30-degree manual bombing task was chosen as a basic maneuver that is learned by all attack pilots and one that offers a substantial learning challenge.

PERFORMANCE MEASUREMENT

Performance measurement was regarded as an important design issue for this experiment. Gross measures, such as bomb miss distances, survival rates, and subjective ratings have been used in some air-to-ground research (Gray & Fuller, 1977; Hughes et al., 1979; Kellogg, Prather & Castore, 1981), presumably because they have strong face validity, and because they are relatively easy to collect. Only objective measures were considered in this experiment because they have greater potential for reliability and power.

The two objective performance measurement approaches that have been developed in flight research can be characterized as norm-referenced and criterion-referenced (Ciavarelli, Williams, & Brictson, 1981). The first is based on summary measures of empirical data, and the second on an assumed ideal performance. The norm-referenced approach was discarded when it became apparent that performances of experienced pilots differed widely

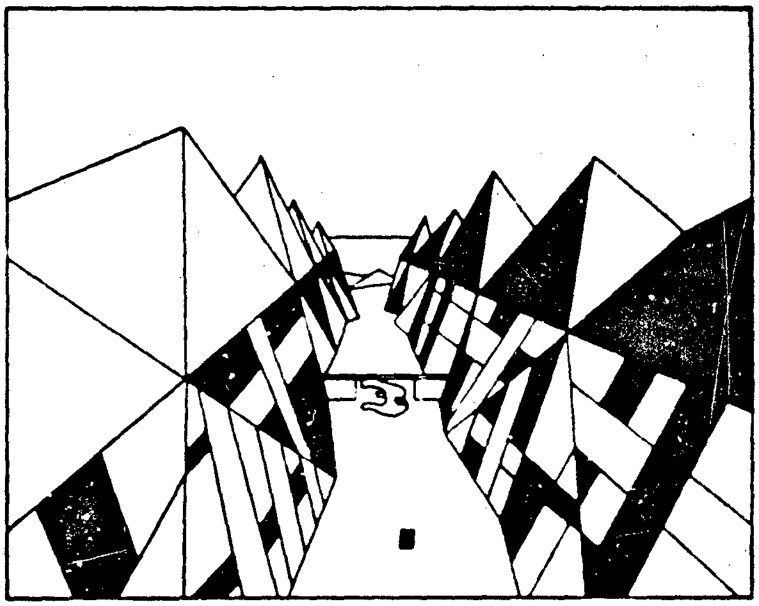


Figure 3. Black and White Representation of the Multicolored High-detail River Valley Scene (RV) Used in Transfer. The square in the center foreground represents a white cube that was the designated target.

on some segments of the task. Criterion-referenced performance measures were established for several segments of the task, based on an analysis of the formal task requirements.

Although there was some difficulty in precisely determining a criterion performance from the task requirements, the one that was determined resulted in a coherent task that could be executed accurately and seemed to be a sensible compromise of ambiguities in the definition of the task. It was assumed that the purposes of this experiment would be served if an explicit and reasonable criterion were established and clearly described to the pilots (Roscoe & Childs, 1980). This would enhance the likelihood that all subject pilots would be attempting to learn the same thing, and should stabilize the data to some extent. Validity was established within the experiment by using only those measures that reflected significant learning.

STATISTICAL POWER

A second major design issue is that of statistical power, defined as the probability that an experiment will show a real effect to be statistically significant. This issue has been largely ignored in applied psychological research, but Cohen (1977) has outlined the rationale and procedures so well that continued neglect seems inexcusable. Applied research, and in particular transfer-of-training research in aviation, is so costly and difficult that researchers should ensure at the outset that their experiments will be sufficiently powerful to demonstrate important effects. Unfortunately the difficulty and expense of transfer-of-training experiments in aviation appear to have encouraged the use of fewer subjects and thereby have led to experiments with very low power (Waag, 1981).

A discussion of power is included as Appendix A. A power criterion of 0.95 for a medium effect (as described in Appendix A), with a significance criterion of 0.05, was established for this experiment. Calculations based on this power criterion (also shown in Appendix A) indicated that 32 experimental subjects would be sufficient if an effective covariate was available for the data analysis. A video game has previously been used as an effective covariate in carrier landing research (Lintern & Kennedy, 1982). That game had been selected because of a demonstrated association between it and compensatory tracking (Kennedy, Bittner, & Jones, 1981). Although bombing is primarily a pursuit task, the selected video game is also a pursuit task and was thought to offer potential as a covariate for this research.

Power analyses were planned for the data that were to be collected in this experiment. In the event that power could not be improved to 0.95 for a medium effect with a significance criterion of 0.05, a supplementary significance level would be established at 0.10. Null hypotheses that were rejected at the

0.10 level but not at the 0.05 level would be regarded as provisional and would be verified in later research if they could not be verified by other data from this experiment.

PROBE METHODOLODY

To explore the immediate effects of augmented feedback, probe trials without augmented feedback were interspersed with augmented feedback trials to test carryover from augmented to nonaugmented conditions.

Hughes, Lintern, Wightman, Brooks, and Singleton (1982) have argued that probe methodology offers several advantages for early investigation of new training strategies. In particular, it avoids the need to establish an optimum training time for the augmented feedback condition, a process that may be expensive and may produce a result that is incompatible with satisfactory exploration of other issues. Hughes et al. (1982), established that probe methodology could identify learning effects and learning differences due to different experimental treatments. Thus, it was considered here as an appropriate supplement to the standard transfer-of-training design.

SECTION II

SUMMARY OF RESEARCH OBJECTIVES

1. To examine the effects of scene content on training performance and on differential transfer*.

2. To assess performance and differential transfer effects of training with only one or with two visual scenes.

3. To assess the differential transfer effects of adaptive supplementary visual guidance in training.

4. To examine interactions between prior air-to-ground bombing experience and other experimental manipulations.

---

*The term <u>differential transfer</u> is used to indicate the difference in performance on a criterion or test condition between groups that have had prior training in different conditions. This term should be distinguished from <u>transfer</u> which is used to indicate the difference in performance on a criterion or test condition between groups, only one of which has had prior training.

SECTION III

METHOD

Sixteen Air Force and 16 Navy pilots were taught a 30-degree cone bombing pattern (Appendix B) for manual delivery of a 25 lb. Mark 76 bomb. The Air Force pilots had qualified from their undergraduate T-37 and T-38 training but had no prior air-to-ground attack experience. The Navy pilots had graduated from TA-4 training and had made 60 to 100 weapon practice deliveries predominantly from a 30-degree dive.

APPARATUS

The Visual Technology Research Simulator (VTRS) is a fully instrumented T-2C Navy, jet trainer cockpit on a six-degree-of-freedom synergistic motion platform with a g-seat, a wide-angle visual system, an A-4 weapon control panel, and an Experimenter/Operator Control Station. The motion system and g-seat were not activated in this experiment.

The visual system has a 1025-line raster and can project colored images onto the inside of a 10-ft radius sphere. The field-of-view extends 50 degrees above to 30 degrees below the pilot's eye level, and 120 degrees left to 40 degrees right of straight ahead from the cockpit. A 1025-line target projector that is available to enhance the target area was not used.

Three visual scenes were used (Figures 1, 2 and 3). One, referred to as the Landscape (LS), was of a flat terrain with multicolored fields, roads, towns, and isolated buildings. One of the isolated buildings was selected as the target. Another scene, referred to as the Grid Pattern (GP), had a white grid laid over a green background. The target was a white cube at the center of one of the grid squares. The remaining scene, known as the River Valley (RV), was of a river paralleled by mountain ranges on either side and approximately a half mile apart. The mountain peaks extended to an altitude of approximately 4000 feet. Two bridges, a small town, and some vessels were located along the river between the mountain ranges. The first two scenes were used in both the training and transfer phases, while the third was used only in transfer. Scene brightness readings are shown in Table 1.

TABLE 1. SCENE BRIGHTNESS READINGS*

Visual Scene	Target		Surround	
Landscape	Walls Roof	Ø.2Ø fL Ø.15 fL	Green Field	Ø.13 fL
Grid Pattern	Cube	1.74 fL	Green Background	Ø.3Ø fL
River Valley	Cube	1.55 fL	Blue River	Ø.29 fL

*All measures taken at an altitude of 6ØØØ feet in the dive to the target from the pilot's eye position.

Average delay between control inputs and generation of the corresponding visual scene was approximately 117 msec (calculation of new aircraft coordinates required 5Ø msec, while calculation of the visual scene corresponding to the viewpoint from the new aircraft coordinates required approximately 5Ø msec, and generation of the new scene required 17 msec). An updated visual scene was displayed every 33 msec.

The bombsight was simulated with a slide projector and a transparency of the mil rings and markings found in some conventional bombsights (Figure 4). The transparency image was projected onto the screen in front of the cockpit and could be calibrated for various weapons. This bomb-sight simulation differed in some obvious ways from a real bomb sight, which is mounted above the instrument panel and requires that the pilot adjust his eye position so that he can look through it.

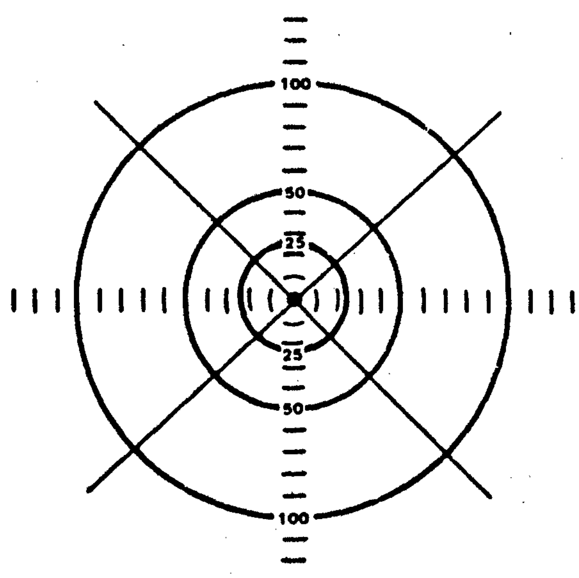


Figure 4. Bomb Sight

The simulator could be set at the optimum position, state, and altitude for weapon release to check the bomb-sight calibration. The scoring algorithm could also be tested at this initial point with altitude and ground position frozen but with attitude and aircraft state unfrozen. Bombs released in this condition typically scored within 20 feet of the target (within 40 feet is considered excellent in a normal bombing run). Some score variation was to be expected with this test because of the difficulty of maintaining precise heading and pitch.

TASK DEFINITION

Details of the bombing pattern are outlined in Appendix B. Initial engine power was set at 96%, with the simulator in straight-and-level flight at 250 knots airspeed, 1000 feet behind the abeam-target position, and at an altitude of 8000 feet. The elevator trim was set for a 30-degree dive at 350 knots airspeed so that, on release from freeze, pilots had to apply back pressure on the control stick to maintain altitude. The heading of the run-in line to the target was displaced from the initial heading by 180 degrees.

Pilots were instructed to fly to the abeam-target position, which was at a ground distance of 12,938 feet from target, and then to enter a 30-degree banked turn to the left while maintaining altitude and radial distance from the target. This constant-altitude portion of the task is often referred to as the cone segment because student pilots are instructed to visualize it as a portion of the base of an inverted cone that has its apex at the target. At an angle of 30 degrees from the run-in line (subtended at the target by ground-plane reference), bank was to be increased to 45 degrees. Altitude was to be maintained at 8000 feet, while distance from the target would decrease. At an angle of 10 degrees from the run-in line, pilots were to reduce power to 86%, roll to a bank angle of 120 degrees, and to pull the 120-mil ring of the bombsight towards the target. The nose of the aircraft was to be allowed to drop to approximately 30 degrees below the horizon. As the 120-mil ring of the bombsight approached the target, the simulated aircraft was to be rolled upright on the run-in line and heading in a 30-degree dive.

A straight-path tracking method was used in which the aircraft is held in a constant dive and the pipper of the bombsight is caused to track towards the target. A curved dive path, in which the aircraft is rolled out with the 40-mil ring on the target, is more usual. With this method the pilot holds the 40-mil ring on the target to an altitude of approximately 4500 feet and then causes the pipper to track towards the target. Thus, the dive angle increases slightly during the first portion of the dive, and the flight path curves downward. Preexperimental pilots commented that control forces in the VTRS were so much higher than was normal for the T-2 aircraft that

curvilinear tracking was difficult. Thus, pilots were instructed to use the straight-path tracking method.

In the optimum dive to the target, airspeed would increase from 250 knots to 350 knots with a dive angle maintained at a constant 30 degrees. The designated release altitude was 3000 feet so that the pilot's goal was to have an airspeed of 350 knots, a dive angle of 30 degrees, and the desired heading along the run-in line as the pipper tracked through the target and as the simulated aircraft passed through 3000 feet. Pilots were advised that, due to altimeter lags, an altimeter reading of 3450 feet corresponded to an actual altitude of 3000 feet for a 30-degree dive (these lags occur in the aircraft as a result of static-pressure operation of the altimeter). The bomb, if released at that point, would score a close hit.

Pilots were also advised that if at an altitude of approximately 4500 feet they could ascertain that one or more of the release parameters would be in error at the release point, specific compensatory adjustments were possible. For example, if the pipper appeared to be too close to the target, it would most likely pass through the target before the designated altitude was reached. However, an early release at higher than designated altitude would compensate for the error and still permit a close hit. A table of compensatory corrections was provided.

EXPERIMENTAL DESIGN

A transfer-of-training design with 80 training and 30 transfer trials was used. One-quarter (8) of the pilot subjects were trained with the Landscape and another quarter (8) were trained with the Grid Pattern. The remaining half (16) were trained with both in a sequence in which a subject would receive three trials with one scene, three with the other, and then another one with each. This series of eight trials was repeated throughout the 80 training trials. The starting heading was 180 degrees and the run-in heading 360 degrees for all training trials.

Augmented feedback was fully crossed with the visual scene factor so that half the subjects practiced with supplementary visual cues and half did not. Augmented-feedback locations were defined on the optimum flight path and at 90-, 60-, 30-, and 10-degree offsets from the run-in line (subtended at the target by the ground-plane reference) in the constant-altitude (cone) portion of the pattern and at altitudes of 6000, 4500 and 3000 feet in the dive. The supplementary visual cues were yellow cubes suspended in the air. They were placed 500 feet ahead of the defined augmented-feedback points so that their offset angle could not become so extreme before a defined augmented-feedback point was passed that they could not be seen. Only one of the supplementary visual cues was active at any time, and control

would pass to the next as one was passed. The augmented feedback was available only during training and only in the first six trials of each eight-trial series. The final two trials of each eight-trial series were considered as probe trials and were compared to similarly placed trials of the nonaugmented groups to assess the effect of training with augmented feedback.

The augmented feedback was adaptive in that the supplementary visual cues appeared only when error limits were exceeded. In the constant altitude segment of the task supplementary cues would switch on only if the altitude error exceeded 200 feet. They would switch on during the dive only if the altitude error exceeded 100 feet. The two levels of subject experience were fully crossed with the display and augmented feedback factors.

In transfer all subjects flew with the two scenes used in training plus the River Valley. Run-in headings for the Landscape and Grid Pattern were 360 and 225 degrees. The two run-in headings for the River Valley were 180 and 360 degrees. The two River Valley tasks provided noticeably different conditions in that the 360-degree run-in required pilots to circle around the end of a mountain range, while the 180-degree run-in required pilots to fly over one of the mountain ranges to reach the run-in line. Initial headings were displaced 180 degrees from run-in headings. The desired flight path was identical to that used in training, but augmented feedback was not provided.

Thirty trials of a psychomotor video game called Air Combat Maneuvering (ACM) were administered to obtain scores that could be used as a covariate in the analysis. ACM is an Atari video game that has previously shown an association with simulated carrier landing performance (Lintern & Kennedy, 1982).

PROCEDURE

Pilots arrived at the VTRS facility either two or three at a time. They were briefed on the experiment and on the procedures for the task. An instruction pamphlet was prepared for this purpose (Berry & Lintern, 1982) and is available from the VTRS facility on request. Pilots were shown the cockpit of the simulator, and its important features were identified and described. They were explicitly requested not to watch the visual displays as others were flying the simulator or being tested on the video tasks. (Experimenters monitored this throughout the experiment and achieved a high degree of compliance).

Ten trials of ACM were administered before the first simulator flight. Simulator training trials were conducted over five 16-trial blocks. Pilots were instructed to drop a bomb at

every trial even if they were badly out of the specified limits. They were advised of their miss distance (e.g., 155 feet at four o'clock), and of their release parameters (airspeed, altitude, and dive angle) after each trial, including transfer trials. There were rests of at least 1.5 hours between sessions. Twenty further ACM trials, three 22.5-minute blocks of Breakout (another Atari video game), and several short pencil-and-paper tests were administered between simulator sessions. (These data were gathered for other purposes and are not reported here.)

The 30 transfer trials followed the final training session (either in the afternoon, if the last training block were in the morning, or next morning, if the last training block were in the afternoon). The break between transfer sessions was at least 1.5 hours. Pilots were given a debriefing interview and they filled out a short biographical questionnaire during this break. The biographical questionnaire was administered primarily to check possibilities for further research. Responses to the questionnaire are not considered in this report.

PERFORMANCE MEASUREMENT

The optimum flight path for the constant-altitude portion of the pattern and for the dive to the target was calculated as shown in Appendix B. This calculation was based as closely as possible on descriptions of the task obtained from flight training instructions for the TA-4J (Naval Air Training Command, 1980). An appropriate description of air-to-ground bombing for the T-2 is not available because that aircraft is not used for air-to-ground attack. Some aspects of the TA-4J task (e.g., bank angles and airspeeds) were modified to be consistent with the performance envelope of the T-2 simulation.

Measures were selected to represent dimensions of performance that should reflect varying levels of skill. Validation of the measures was accomplished by examining learning trends in early trials.

SECTION IV

RESULTS

EVALUATING THE MEASURES

Performance measures considered for analysis are shown in Table 2 together with levels of statistical significance and values of Eta squared (the variance accounted for by an effect) for those measures during the training trials. Error in the cone pattern was examined. The learning trends are shown in Figure 5. With the exception of RMS lateral error in the cone segment of the task, all learning effects were significant and substantial. This evidence was used to validate the use of these measures to examine differential quality of performance. The lack of any significant learning effect with the measure of RMS lateral error in the cone segment cast some doubt on its validity as a measure of the quality of performance. The only other means of validating this measure was to look for differences between levels of pilot experience. As there were no significant differences, this measure was not considered for further analysis.

Table 3 shows calculations between ACM and the performance measures (with effects of factors partialled out). Only the correlation with RMS pitch error in the final approach was sufficiently large to warrant the use of ACM as a covariate. The fact that this association was negative would not necessarily mean that ACM should not be used as a covariate. However, the pattern of small correlations, with two showing a negative association, cast doubt on the validity of using ACM as a covariate. Thus, the variances were analyzed without taking covariate performance into account.

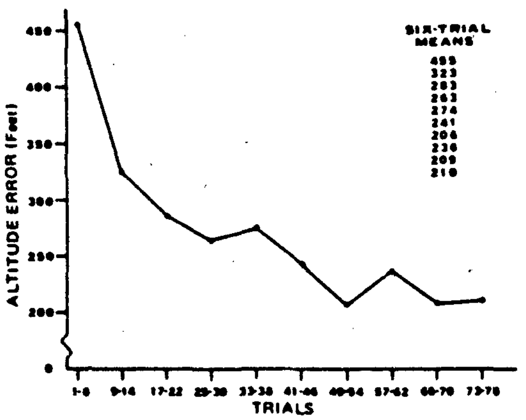
A summary of the statistical power analyses is shown in Table 4. Miss distances of 40, 70, and 100 feet were considered small, medium, and large, respectively (as noted in Appendix A), and were transformed into altitude, pitch, and lateral errors in the dive by finding the magnitudes of error at bomb release that would result in bomb miss distances of 40, 70, and 100 feet. The values for the final approach measures are low but for the remaining measures are excellent. The tabled power values are reasonable estimates for main effects and for interactions. However, the power of the _post hoc_ paired comparisons is lower, especially for tables with large numbers of cells.

TABLE 2. STATISTICAL SIGNIFICANCE AND SIZE OF LEARNING EFFECT IN TRAINING TRIALS FOR SELECTED PERFORMANCE MEASURES

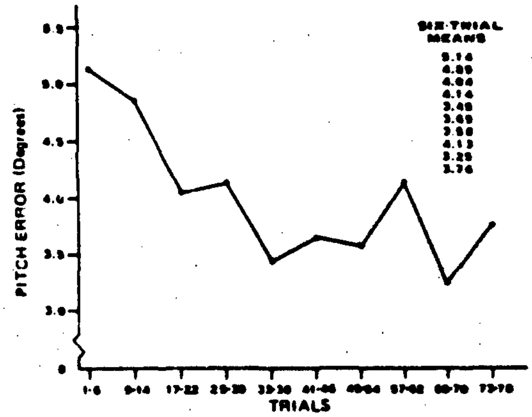
	p	eta
Target Approach		
RMS Altitude Error	0.001	.22
RMS Pitch Error	0.001	.12
RMS Lateral Error	0.005	.09
Bomb Impact		
Absolute Longitudinal Miss Distance	0.001	.27
Absolute Lateral Miss Distance	0.001	.16
Cone Segment		
RMS Altitude Error	0.001	.68
RMS Lateral Error	0.171	-

TABLE 3. CORRELATIONS BETWEEN AIR COMBAT MANEUVERING PERFORMANCE AND TRANSFER PERFORMANCE (SIGNS REFLECTED SO THAT POSITIVE CORRELATIONS SHOW A POSITIVE ASSOCIATION)

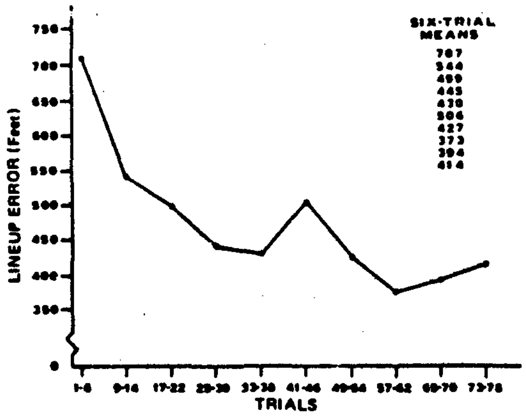
Final Approach	
RMS Altitude Error	-.18
RMS Pitch Error	-.40
RMS Lateral Error	.09
Bomb Impact	
Absolute Longitudinal Miss Distance	01
Absolute Lateral Miss Distance	.03
Cone Pattern Segment	
RMS Altitude Error	.08



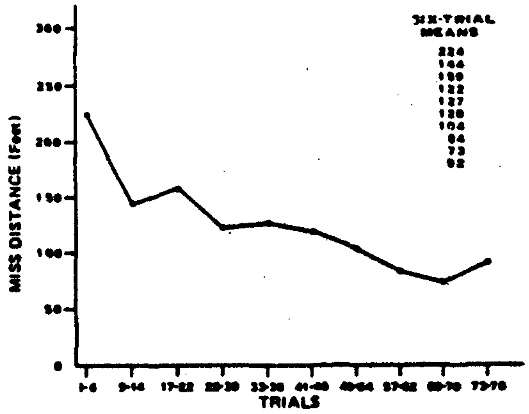
(a) RMS Altitude Error in the Dive.



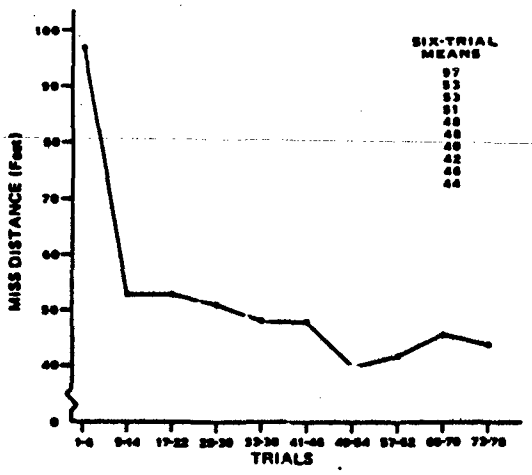
(b) RMS Pitch Error in the Dive.



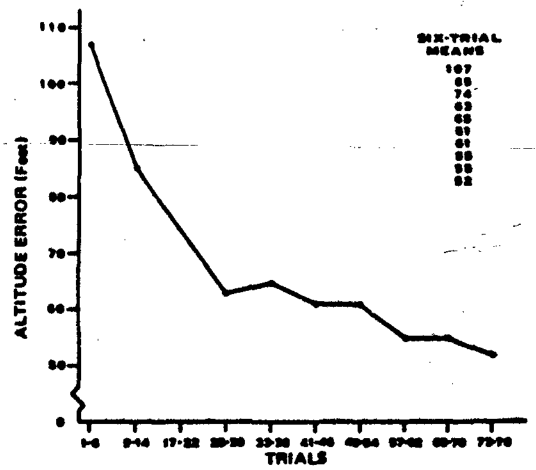
(c) RMS Lineup Error in the Dive.



(d) Absolute Longitudinal Bomb Miss Distance.



(e) Absolute Lateral Bomb Miss Distance.



(f) RMS Altitude Error in the Cone Segment.

Figure 5. Learning Trends in Training Trials

TABLE 4. POWER COEFFICIENTS FOR TRANSFER DATA; $p = .05$ (.10)

Effect	Small		Medium		Large	
	Effect Size	Power	Effect Size	Power	Effect Size	Power
Final Approach Measures						
RMS Altitude Error	44 ft	.17 (.27)	77 ft	.45 (.59)	111 ft	.78 (.88)
RMS Pitch Error	.42 deg	.08 (.15)	.73 deg	.16 (.26)	1.04 deg	.29 (.42)
RMS Lateral Error	40 ft	.06 (.11)	70 ft	.07 (.14)	100 ft	.10 (.17)
Bomb Impact Measures						
Absolute Longitudinal Miss Distance	40 ft	.58 (.72)	70 ft	.98 (.99)	100 ft	.99+(.99+)
Absolute Lateral Miss Distance	40 ft	.99 (.99)	70 ft	.99+(.99+)	100 ft	.99+(.99+)
Cone Pattern Measure						
RMS Altitude Error	25 ft	.76 (.86)	50 ft	.99+(.99+)	100 ft	.99+(.99+)

PRESENTATION OF DATA

In reviewing the results in the tables that follow, readers should recall the varying relationships of the experimental factors to training, probe, and transfer trials. The training trials represent a situation in which the pilot subjects flew repeated trials in the simulator on one of the experimental conditions, interrupted periodically by probe trials. In the transition from training to probe trials, only the relationship of the Augmented-Feedback factor changed. Augmented Feedback was not presented during probe or transfer trials. In transfer, all experimental subjects flew an identical series of scenes and run-in headings, only some of which had been used in training and probe trials. The training and transfer results are presented in Tables 5 through 16. Results for the probe trials and their relationships to the training and transfer results are summarized in Table 17, and some important learning trends across probe trials are presented in Table 18.

The considerable data that were available for presentation forced further selection. All main effects and significant ($p < .05$) interactions for training and transfer trials are presented in Tables 5 through 16. Additionally, two-way interactions are presented for final approach measures if their probability of statistical significance is between 0.05 and 0.10. This was done primarily because of the low power of the approach measures. Results that were significant only at this relaxed level were not examined further with paired comparisons and were given little weight in the discussion unless other data supported their trend. ANOVA tables for the six dependent measures are shown in Appendix C.

Deviations of observed from expected cell means were used as an aid to interpretation of two- and three-way interactions not involving repeated measures. Large deviations (both positive and negative) indicate departures from overall trends. Critical differences based on t tests are presented for significant main effects involving more than two cells. These values indicate magnitudes of difference between pairs of means that can be considered statistically significant. They are used to ensure that differences between pairs of means that appear large are not emphasized in the interpretation of the data if they do not approach statistical significance. Cell deviations of significant interactions were tested for statistical significance with an F test described by Jones (unpublished). Where learning trends are to be assessed, critical differences based on t tests are used to estimate significant improvement.

TABLE 5. SUMMARY OF IMPORTANT BOMBING-RUN PERFORMANCE EFFECTS DURING TRAINING TRIALS, BASED ON GEOMETRIC MEANS OF RMS ALTITUDE ERRORS IN FEET

Source of Variance	Cell Means	F	p	η^2
Scene Type (ST)		3.19	0.063	0.15
Landscape (LS)	215	Altitude control was better with the LS scene only than with the GP or with both during training.		
Grid Pattern (GP)	274			
Both (LS=280, GP=287)	283			
Augmented Feedback (FB)		4.82	0.040	0.11
FB On	237	Augmented feedback facilitated altitude control during training.		
FB Off	291			
Prior Experience (PE)		3.79	0.066	0.09
No PE	281	Pilots with some prior bombing experience performed better.		
Some PE	246			
Augmented FB X Prior Experience		4.03	0.058	0.09
No Prior Experience		The significant main effect of augmented feedback was entirely due to the enhancing effect of of augmented feedback on the performance of pilots with no prior bombing experience.		
FB On	230 (-27)+			
FB Off	342 (+27)			
Some Prior Experience				
FB On	243 (+27)			
FB Off	247 (-27)			

+In this table and those that follow, values in parentheses are deviations of the observed cell means from those expected based on main factor effects had there been no interaction.

TABLE 6. SUMMARY OF IMPORTANT BOMBING-RUN PERFORMANCE EFFECTS DURING TRANSFER TRIALS, BASED ON GEOMETRIC MEANS OF RMS ALTITUDE ERRORS IN FEET

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		3.51	0.049	0.18	76
Landscape (LS)	19[illegible]	Training with the LS scene only facilitated transfer more than training with the GP or both.			
Grid Pattern (GP)	284				
Both (LS & GP)	276				
Augmented Feedback (FB)		4.54	0.046	0.12	...
FB On	224	Augmented feedback during training facilitated transfer.			
FB Off	293				
Prior Experience (PE)		2.65	0.119	0.07	
No PE	272	The difference in favor of the more experienced pilots was not statistically significant.			
Some PE	241				
Scene Type during Transfer		3.74	0.004	0.08	50
Landscape (LS)		The familiar 360-degree run-in resulted in better performance with the Landscape scene than did the unfamiliar 225-degree run-in. No such trend was evident with the Grid Pattern. Performances with the unfamiliar River Valley were better with the 180-degree run-in (pilots flew over a mountain range to the run-in line) than within the 360-degree run-in (pilots flew around a mountain range).			
360-degree Run-in	210				
225-degree Run-in	271				
Grid Pattern (G					
360-degree Run-in	254				
225-degree Run-in	254				
River Valley (RV)					
360-degree Run-in	310				
180-degree Run-in	247				

TABLE 6. (Altitude Errors, Bombing Run, Transfer Trials continued)

Source of Variance	Cell Means	F	p	η^2
Training ST X Transfer ST		2.30	0.018	0.10
Landscape during Training				
LS-360 during Transfer	163 (+ 5)			
LS-225	214 (- 7)			
GP-360	180 (-24)			
GP-225	131 (-84)**			
RV-360	284 (+29)			
RV-180	268 (+76)**			
Grid Pattern during Training				
LS-360 during Transfer	218 (-20)			
LS-225	307 (+08)			
GP-360	278 (-06)			
GP-225	317 (+22)			
RV-360	349 (+14)			
RV-180	225 (-17)			
Both (LS & GP) during Training				
LS-360 during Transfer	235 (+07)			
LS-225	287 (-02)			
GP-360	288 (+14)			
GP-225	316 (+31)*			
RV-360	304 (-21)			
RV-180	233 (-29)			

The main effect trend of superior performance following Landscape training was not evident in transfer to the unfamiliar River Valley scene with the 180-degree run-in but was accentuated on transfer to the Grid Pattern with the new 225-degree run-in. Transfer performances on the Grid Pattern scene showed evidence of asymmetrical transfer, in that those trained on the Landscape scene outperformed those trained on the Grid Pattern or on both scenes. Tukey tests of the relevant paired comparisons were significant (critical differences = 163 for q (18,100) = 4.93) in the context of transfer to the 225-degree run-in.

In this table and those that follow, * indicates $p < 0.05$; ** indicates $p < 0.01$.

There are insufficient degrees of freedom to test all cells. In this interaction, and those that follow, some deviations that may be significant must remain untested.

TABLE 6. (Altitude Errors, Bombing Run, Transfer Trials continued)

Source of Variance	Cell Means	F	p	η^2
Augmented FB X Transfer ST		5.44	0.001	0.12
FB On during Training				
LS-360 during Transfer	147 (+44)*	The main effect of superior performance following augmented feedback training was absent (or possibly even reversed) on transfer to the Landscape scene and the unfamiliar 225-degree run-in.		
LS-225	292 (+53)			
GP-360	201 (-27)			
GP-225	227 (+05)			
RV-360	282 (-28)			
RV-180	227 (-38)			
FB Off during Training				
LS-360 during Transfer	301 (+44)			
LS-225	251 (-53)**			
GP-360	320 (+27)			
GP-225	283 (-05)			
RV-360	340 (+28)			
RV-180	269 (+38)			
Transfer ST X Prior Experience		2.91	0.17	0.06
No Prior Experience				
LS-360 during Transfer	191 (-39)	The main effects of good transfer performance with the Landscape scene and the 360-degree run-in and poor transfer performance with the River Valley and the 360-degree run-in are restricted to pilots with no prior bombing experience.		
LS-225	287 (-03)			
GP-360	253 (-20)			
GP-225	260 (-13)			
RV-360	390 (+53)			
RV-180	289 (+20)			
Some Prior Experience				
LS-360 during Transfer	232 (+39)*			
LS-225	255 (+03)			
GP-360	255 (+20)			
GP-225	248 (+13)			
RV-360	246 (-53)**			
RV-180	211 (-20)			

TABLE 6. (Altitude Errors, Bombing Run, Transfer Trials continued)

Source of Variance	Cell Means	F	p	η^2	t
Transfer ST X Transfer Trials		1.80	0.019	0.06	109

	Transfer Trials				
Transfer Scene Type	01-06	07-12	13-18	19-24	25-30
Landscape (LS)					
360-degree Run-in	213	222	174	255	195
225-degree Run-in	481	197	243	259	244
Grid Pattern (GP)					
360-degree Run-in	185	250	233	319	306
225-degree Run-in	291	215	215	251	311
River Valley (RV)					
360-degree Run-in	343	322	347	273	272
180-degree Run-in	279	262	254	242	203

This interaction is due largely to improving performance on the Landscape scene with the new 225-degree run-in, and deteriorating performance on the Grid Pattern with the familiar 360-degree run-in. The learning trend shown with the Landscape scene and the 225-degree run-in heading is due entirely to the disproportionately poor altitude control on the first block of transfer trials.

TABLE 6. (Altitude Errors, Bombing Run, Transfer Trials continued)

Source of Variance	Cell Mean	F	p	η^2
Augmented FB X Transfer ST X PE		2.34	0.047	0.05
No Prior Experience (PE)				
FB On during Training				
LS-360 during Transfer	136 (+05)			
LS-225	283 (-30)*			
GP-360	197 (-11)			
GP-225	243 (+04)			
RV-360	361 (-09)			
RV-180	322 (+40)**			
FB Off during Training				
LS-360 during Transfer	269 (-05)			
LS-225	291 (+30)			
GP-360	324 (+11)			
GP-225	278 (-04)			
RV-360	421 (+09)			
RV-180	258 (-41)			
Some Prior Experience				
FB On during Training				
LS-360 during Transfer	154 (-05)			
LS-225	301 (+30)			
GP-360	206 (+11)			
GP-225	212 (-04)*			
RV-360	220 (+09)			
RV-180	154 (-40)			
FB Off during Training				
LS-360 during Transfer	337 (+06)			
LS-225	217 (-30)*			
GP-360	316 (-11)			
GP-225	290 (+04)			
RV-360	274 (-08)			
RV-180	280 (+41)*			

The main effect of superior performance following augmented feedback training was clearly reversed in the transfer of the more experienced pilots to the Landscape scene and 225-degree run-in, and also in the transfer of the inexperienced pilots to the River Valley with the 180-degree run-in.

TABLE 7. SUMMARY OF IMPORTANT BOMBING-RUN PERFORMANCE EFFECTS DURING TRAINING TRIALS, BASED ON GEOMETRIC MEANS OF RMS PITCH ERRORS IN DEGREES

Source of Variance	Cell Means	F	p	η^2	t
Scene Type (ST)		13.29	0.001	0.27	1.07
Landscape (LS)	2.48	Pitch control during training was better with the LS scene only than with the GP or with both.			
Grid Pattern (GP)	4.77				
Both (LS = 4.37, GP = 4.75)	4.51				
Augmented Feedback (FB)		5.60	0.028	0.06	
FB On	3.47	Augmented feedback facilitated pitch control during training.			
FB Off	4.52				
Pilot Experience (PE)		0.55	0.465	0.01	
No PE	4.12	Prior experience had no significant effect by this measure.			
Some PE	3.82				
Augmented Feedback X Scene Type		6.05	0.009	0.12	
Landscape during Training					
FB On	2.88 (+1.03)	The significant main effect trend of better performance with augmented feedback was not evident when only the Landscape was used in training.			
FB Off	2.13 (-1.04)				
Grid Pattern during Training					
FB On	3.36 (-0.97)**				
FB Off	6.62 (+0.97)				
Both (LS & GP) during Training					
FB On	3.85 (-0.04)				
FB Off	5.25 (+0.04)				

TABLE 7. (Pitch Errors, Bombing Run, Training Trials continued)

Source of Variance	Cell Means	F	p	η^2	t
Augmented FB X Prior Experience		22.58	0.001	0.23	
No Prior Experience					
FB On	2.88 (-.88)	The facilitating effect of augmented feedback in training was restricted to those pilots with no prior bombing experience.			
FB Off	5.76 (+.88)**				
Some Prior Experience					
FB On	4.14 (+.88)				
FB Off	3.52 (-.88)				
Augmented FB X Training Trials		2.08	0.034	0.06	1.05

Training Trials

	01-06	09-15	18-22	25-30	33-38	41-46	49-54	57-62	65-70	73-78
Augmented FB On	5.22	5.08	3.70	3.80	2.96	2.92	2.99	3.43	2.45	2.89
Augmented FB Off	5.07	4.64	4.40	4.50	3.98	4.51	4.25	4.96	4.22	4.83

This significant interaction is due to the rapid early improvement in training performances when augmented feedback was provided; without such FB performances improved little during training trials.

TABLE 8. SUMMARY OF IMPORTANT BOMBING-RUN EFFECTS DURING TRANSFER TRIALS, BASED ON GEOMETRIC MEANS OF RMS PITCH ERRORS IN DEGREES

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		7.14	0.005	0.26	1.47
Landscape (LS)	1.88	Training with the LS scene only facilitated transfer more than training with GP or with both.			
Grid Pattern (GP)	4.02				
Both (LS & GP)	3.91				
Augmented Feedback (FB)		4.82	0.040	0.11	
FB On	2.71	Augmented feedback during training facilitated pitch-control transfer.			
FB Off	4.05				
Prior Experience (PE)		0.60	0.447	0.01	
No PE	3.48	Prior experience had no significant effect on transfer performance.			
Some PE	3.17				
Scene Type during Transfer		7.27	0.001	0.17	0.42
Landscape (LS)					
360-degree Run-in	3.37	Pitch control during transfer trials was superior with the River Valley scene and inferior with the Grid Pattern, regardless of the run-in heading.			
225-degree Run-in	3.25				
Grid Pattern (GP)					
360-degree Run-in	3.98				
225-degree Run-in	3.82				
River Valley (RV)					
360-degree Run-in	2.78				
180-degree Run-in	2.88				

TABLE 8. (Pitch Errors, Bombing Run, Transfer Trials continued)

Source of Variance	1 Means	F	p	η^2
Training ST X Transfer ST		2.33	0.016	0.11
Landscape during Training				
LS-360 during Transfer	1.62 (-.34)**			
LS-225	1.94 (+.18)			
GP-360	2.24 (-.31)			
GP-225	1.72 (-.76)**			
RV-360	1.88 (+.64)**			
RV-180	1.96 (+.60)**			
Grid Pattern during Training				
LS-360 during Transfer	4.51 (+.37)*			
LS-225	3.85 (-.09)			
GP-360	4.86 (+.13)			
GP-225	4.58 (-.08)			
RV-360	3.24 (-.18)			
RV-180	3.36 (-.18)			
Both (LS & GP) during Training				
LS-360 during Transfer	4.01 (-.02)			
LS-225	3.78 (-.05)			
GP-360	4.70 (-.08)			
GP-225	4.96 (+.41)**			
RV-360	3.07 (-.24)			
RV-180	3.21 (-.22)			

The significant main effect of better performance on transfer to the River Valley scene was evident following training on the Grid Pattern or both scenes but not following training on the Landscape only. Transfer to the Grid Pattern showed evidence of asymmetrical transfer. Tukey tests of the four relevant paired comparisons showed that pilots trained on the Landscape outperformed those trained on the Grid Pattern or on both scenes (critical difference = 1.36 degrees for q(18, 100) = 4.93).

Source of Variance	1 Means	F	p	η^2
Augmented FB X Prior Experience		6.92	0.016	0.12
No Prior Experience				
FB On	2.33 (-.60)**			
FB Off	5.03 (+.60)			
Some Prior Experience				
FB On	3.12 (+.60)			
FB Off	3.22 (-.60)			

The significant effect of augmented feedback during training on pitch-control skill was restricted to pilots with no prior bombing experience.

TABLE 9. SUMMARY OF IMPORTANT BOMBING-RUN PERFORMANCE EFFECTS DURING TRAINING TRIALS, BASED ON GEOMETRIC MEANS OF RMS LATERAL ERRORS IN FEET

Source of Variance	Cell Means	F	p	η^2
Scene Type during Training		0.26	0.774	0.01
Landscape (LS)	447	Scene type had no significant effect on lateral steering during training trials.		
Grid Pattern (GP)	421			
Both (LS=465, GP=505)	485			
Augmented Feedback (FB)		2.02	0.171	0.05
FB On	410	Augmented FB had no significant effect on lateral steering.		
FB Off	530			
Prior Experience (PE)		3.35	0.082	0.09
No PE	542	Pilots with some prior bombing experience tended to perform better.		
Some PE	401			
Scene Type X Prior Experience		3.42	0.053	0.19
No Prior Experience				
Landscape (LS)	768 (+160)	The near significant main trend of better performance by the more experienced pilots was accentuated by Landscape training. Pilots with no prior bombing experience did poorly on this scene.		
Grid Pattern (GP)	371 (-130)			
Both (LS & GP)	550 (-016)			
Some Prior Experience				
Landscape (LS)	295 (-155)			
Grid Pattern (GP)	478 (+135)			
Both (LS & GP)	428 (-021)			

TABLE 9. (Lateral Errors, Bombing Run, Training Trials continued)

Source of Variance	F	p	n^2	t
Scene Type X Training Trials	1.93	0.016	0.11	164

	Training Trials									
	01-06	09-14	17-22	25-30	33-38	41-46	49-54	57-62	65-70	73-78
Landscape (LS)	593	499	419	347	423	689	420	503	565	401
Grid Pattern (GP)	495	491	469	445	559	449	443	489	194	324
Both (LS & GP)	924	596	561	504	380	459	422	281	468	477

Learning is evident with all conditions. The interaction is due to disportionately poor lateral control on the first block of trials with both scenes.

Source of Variance	F	p	n^2
Augmented FB X Training Trials	1.86	0.06	0.05

	Training Trials									
	01-06	09-14	17-22	25-30	33-38	41-46	49-54	57-62	65-70	73-78
Augmented FB On	705	518	438	407	377	503	410	305	311	280
Augmented FB Off	710	571	569	485	491	508	443	457	499	612

The pilots who received augmented feedback continued to show improvement in lateral steering control throughout training, whereas those who did not receive augmented feedback showed little if any improvement after the 30th trial and in fact performed quite poorly on the final block of trials.

TABLE 10. SUMMARY OF IMPORTANT BOMBING-RUN PERFORMANCES EFFECTS DURING TRANSFER TRIALS, BASED ON GEOMETRIC MEANS OF RMS LATERAL ERRORS IN FEET

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		0.15	0.860	0.01	
Landscape (LS)	524	Scene type during training had no significant effect on lateral steering during transfer.			
Grid Pattern (GP)	499				
Both (LS & GP)	559				
Augmented Feedback (FB)		0.30	0.587	0.01	
FB On	518	FB had no significant transfer effect by this measure.			
FB Off	568				
Prior Experience (PE)		1.24	0.279	0.05	
No PE	607	PE also had no significant transfer effect by this measure.			
Some PE	484				
Scene Type during Transfer		8.76	0.001	0.19	138
Landscape (LS)					
360-degree Run-in	458 (-101)	For transfer to the Landscape and the Grid Pattern, lateral steering control was poorer with the new 225-degree run-in.			
225-degree Run-in	758 (+199)				
Grid Pattern (GP)					
360-degree Run-in	462 (-097)				
225-degree Run-in	752 (+193)				
River Valley (RV)					
360-degree Run-in	409 (-150)				
180-degree Run-in	514 (-045)				

TABLE 10. (Lateral Errors, Bombing Run, Transfer Trials continued)

Source of Variance	Cell Means	F	p	η^2
Scene Type X Prior Experience		2.76	0.022	0.06
No Prior Experience				
LS-360 during Transfer	527 (+015)	The main effect of disproportionately poor transfer to the Landscape with the 225-degree run-in was restricted to the more experienced pilots.		
LS-225	674 (-138)**			
GP-360	573 (+051)			
GP-225	823 (+018)			
RV-360	527 (+055)			
RV-180	568 (+001)			
Some Prior Experience				
LS-360 during Transfer	397 (-015)			
LS-225	850 (+138)			
GP-360	372 (-051)			
GP-225	688 (-018)			
RV-360	317 (-055)			
RV-180	467 (-001)			

TABLE 11. SUMMARY OF IMPORTANT BOMBING PERFORMANCE EFFECTS DURING TRAINING TRIALS, BASED ON GEOMETRIC MEANS OF ABSOLUTE LONGITUDINAL BOMB MISS DISTANCES IN FEET

Source of Variance	Cell Means	F	p	η^2
Scene Type during Training		0.79	0.469	0.04
Landscape (LS)	114	Scene type had no significant effect on longitudinal bomb miss distance during training.		
Grid Pattern (GP)	132			
Both (LS=114, GP=115)	114			
Augmented Feedback (FB)		0.57	0.460	0.02
FB On	123	FB during training also had no significant effect.		
FB Off	115			
Prior Experience (PE)		5.00	0.037	0.14
No PE	131	More experienced pilots performed better by this measure.		
Some PE	107			

TABLE 12. SUMMARY OF IMPORTANT BOMBING PERFORMANCE EFFECTS DURING TRANSFER TRIALS, BASED ON GEOMETRIC MEANS OF ABSOLUTE LONGITUDINAL BOMB MISS DISTANCES IN FEET

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		1.02	0.380	0.06	
Landscape (LS)	96	Scene type during training had no significant effect on transfer performances.			
Grid Pattern (GP)	98				
Both (LS & GP)	106				
Augmented Feedback (FB)		0.72	0.408	0.02	
FB On	100	FB during training also had no significant effect on transfer.			
FB Off	103				
Prior Experience (PE)		1.92	0.181	0.06	
No PE	104	PE had no significant effect on transfer performance.			
Some PE	99				
Scene Type during Transfer		10.28	0.001	0.25	28
Landscape (LS)					
360-degree run-in	87	Longitudinal bomb miss distances during transfer were consistently greater with the new River Valley scene and the 360-degree run-in (pilots flew around the mountains) than with the other five transfer conditions.			
225-degree run-in	79				
Grid Pattern (GP)					
360-degree run-in	98				
225-degree run-in	81				
River Valley (RV)					
360-degree run-in	186				
180-degree run-in	112				

TABLE 12. (Longitudinal Bomb Miss Distances, Transfer Trials continued)

Source of Variance	Cell Means	F	p	η^2
Augmented FB X Prior Experience		4.40	0.049	0.13
No Prior Experience				
FB On	118 (+15)**	Augmented FB during training facilitated the transfer performances of the pilots with some prior bombing experience and had the opposite effect on the inexperienced pilots.		
FB Off	91 (-15)			
Some Prior Experience				
FB On	84 (-15)			
FB Off	115 (+15)			
Transfer Trials		3.54	0.010	0.11

Transfer Trials

01-06	07-12	13-18	19-24	25-30
126	106	91	102	87

Longitudinal bomb miss distances improved irregularly throughout the transfer phase.

TABLE 13. SUMMARY OF IMPORTANT BOMBING PERFORMANCE EFFECTS DURING TRAINING TRIALS, BASED ON GEOMETRIC MEANS OF ABSOLUTE LATERAL BOMB MISS DISTANCES IN FEET

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		4.22	0.030	0.17	15
Landscape (LS)	41	Lateral bomb miss distances were smallest with the LS scene and largest when pilots used LS & GP.			
Grid Pattern (GP)	47				
Both (LS=63, GP=56)	59				
Augmented Feedback (FB)		1.64	0.215	0.03	
FB On	53	By this measure, FB had no significant effect during training.			
FB Off	49				
Prior Experience		2.77	0.100	0.06	
No PE	54	Prior experience had no significant effect by this measure.			
Some PE	47				

TABLE 14. SUMMARY OF IMPORTANT BOMBING EFFECTS DURING TRANSFER TRIALS, BASED ON GEOMETRIC MEANS OF ABSOLUTE LATERAL BOMB MISS DISTANCES IN FEET

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		0.53	0.594	0.03	
Landscape (LS)	50	Scene type during training had no significant effect on lateral bomb miss distances.			
Grid Pattern (GP)	57				
Both (LS & GP)	57				
Augmented Feedback (FB)		0.02	0.88	0.001	
FB On	57	FB had no significant effect on lateral misses during training.			
FB Off	53				
Prior Experience (PE)		0.35	0.56	0.01	
No PE	54	Prior bombing experience also had no effect by this measure.			
Some PE	56				
Scene Type during Transfer		10.65	0.001	0.27	16
Landscape (LS)					
360-degree run-in	38	Lateral bomb miss distances were greater with the new 225-degree run-in than with the familiar 360-degree run-in for pilots transferring to the Landscape or Grid Pattern scenes. With the new River Valley transfer scene, run-in heading appeared to make no difference.			
225-degree run-in	71				
Grid Pattern (GP)					
360-degree run-in	42				
225-degree run-in	98				
River Valley (RV)					
360-degree run-in	53				
180-degree run-in	47				

TABLE 14. (Lateral Bomb Miss Distances, Transfer Trials continued)

Source of Variance	Cell Means	F	p	η^2	t
Training ST X FB X Transfer Trials		2.58	0.015	0.14	32

	Transfer Trials				
	01-06	07-12	13-18	19-24	25-30
Landscape during Training					
FB On	39	67	61	40	25
FB Off	84	43	44	52	71
Grid Pattern during Training					
FB On	70	64	52	95	58
FB Off	46	53	44	56	47
Both (LS & GP) during Training					
FB On	58	52	52	58	69
FB Off	57	51	45	71	49

This interaction effect, though statistically significant, reveals no systematic relationship.

TABLE 15. SUMMARY OF IMPORTANT PERFORMANCE EFFECTS IN THE CONE SEGMENT DURING TRAINING TRIALS, BASED ON GEOMETRIC MEANS OF RMS ALTITUDE ERRORS IN FEET

Source of Variance	Cell Means	F	p	η^2
Scene Type during Training		7.14	0.758	0.01
Landscape (LS)	62	Scene type had no significant effect on cone-segment altitude control.		
Grid Pattern (GP)	69			
Both (LS & GP)	67			
Augmented Feedback (FB)		0.88	0.360	0.02
FB On	69	FB also had no significant effect by this measure.		
FB Off	64			
Prior Experience (PE)		2.98	0.100	0.08
No PE	73	Some PE tended to facilitate training performance slightly.		
Some PE	60			

TABLE 16. SUMMARY OF IMPORTANT PERFORMANCE EFFECTS IN THE CONE SEGMENT DURING TRANSFER TRIALS, BASED ON GEOMETRIC MEANS OF RMS ALTITUDE ERRORS IN FEET

Source of Variance	Cell Means	F	p	η^2	t
Scene Type during Training		0.70	0.508	0.04	
Landscape (LS)	64	Training scene type had no significant effect on cone-segment altitude control during transfer.			
Grid Pattern (GP)	72				
Both (LS & GP)	62				
Augmented Feedback (FB)		0.16	0.691	0.01	
FB On	64	FB during training had no effect on transfer performance.			
FB Off	65				
Prior Experience (PE)		0.62	0.440	0.02	
No PE	66	PE also had no significant effect by this transfer measure.			
Some PE	63				
Scene Type during Transfer		7.81	0.001	0.21	9
Landscape (LS)					
360-degree run-in	56	Cone-segment altitude control during transfer was disproportionately poor with the Landscape scene and 225-degree run-in heading and with the new River Valley scene and the 180-degree run-in in which pilots had to fly over, rather than around, the mountain range.			
225-degree run-in	71				
Grid Pattern (GP)					
360-degree run-in	59				
225-degree run-in	66				
River Valley (RV)					
360-degree run-in	60				
180-degree run-in	79				
Transfer Trials		8.21	0.001	0.23	8

Transfer Trials				
01-06	07-12	13-18	19-24	25-30
79	66	65	61	55

Cone-segment altitude control improved steadily throughout the transfer phase.

TABLE 17. COMPARISON OF PROBE TRIAL PERFORMANCES ON THE VARIOUS MEASURES WITH THE CORRESPONDING TRAINING AND TRANSFER VALUES FOR PILOTS WHO DID AND DID NOT RECEIVE AUGMENTED FEEDBACK DURING TRAINING

	Training Trials		Probe Trials		Transfer Trials	
Performance Measure	Cell Means	p (η^2)	Cell Means	p (η^2)	Cell Means	p (η^2)
Bomb-Run Altitude						
FB On	237	0.040 (.11)	236	0.151 (.05)	224	0.046 (.12)
FB Off	291		275		293	
Bomb-Run Pitch						
FB On	3.47	0.028 (.06)	3.29	0.016 (.08)	2.71	0.038 (.09)
FB Off	4.52		4.61		4.05	
Lateral Steering						
FB On	410	0.171 (.06)	412	0.651 (.01)	518	0.587 (.01)
FB Off	530		447		568	
Longitudinal Bomb Miss						
FB On	123	0.460 (.02)	124	0.536 (.01)	100	0.408 (.02)
FB Off	115		112		103	
Lateral Bomb Miss						
FB On	53	0.215 (.03)	50	0.664 (.01)	57	0.888 (.01)
FB Off	49		54		53	
Cone-Segment Altitude						
FB On	69	0.360 (.02)	65	0.363 (.02)	64	0.691 (.01)
FB Off	64		60		65	

Only the bomb-run altitude and bomb-run pitch measures show significant effects of augmented feedback on performance. The effects for any dependent measure (whether significant or not) are of similar magnitudes across training, probe, and transfer trials.

TABLE 18. SUMMARY OF LEARNING TRENDS IN PROBE TRIALS FOR DIVE PITCH CONTROL

Source of Variance					
FB X PE X Probe Trials					
	Probe Trials				
	01-04	05-08	09-12	13-16	17-20
No Prior Experience					
FB On	4.68	3.16	2.93	2.26	1.97
FB Off	5.46	6.52	5.90	5.87	6.41
Some Prior Experience					
FB On	5.11	4.13	3.42	3.17	2.99
FB Off	3.92	3.67	2.40	3.49	4.13

There is evidence of improvement only with the augmented feedback groups. (Critical difference = 1.52 degrees for t(80) = 1.99)

TABLE 18. (Learning Trends, Probe Trials, Bomb Run Pitch, continued)

Source of Variance

FB X Training ST X Probe Trials

	Probe Trials				
	01-04	05-08	09-12	13-16	17-20
Training Scene Type					
Landscape					
FB On	2.95	2.53	2.98	1.86	1.55
FB Off	2.84	3.18	1.61	1.09	1.97
Grid Pattern					
FB On	5.82	4.37	2.47	2.25	2.36
FB Off	5.84	6.67	5.92	7.04	8.00
Both					
FB On	5.68	3.91	3.70	3.50	2.05
FB Off	5.19	5.21	4.52	6.50	6.35

Learning is significant only for the augmented-feedback groups that trained with the grid pattern or both scenes. The learning trends evident in the two landscape groups are not significant, but their early trials are significantly better than those of the other groups (critical difference = 2.0 degrees to t(80) = 1.49).

SECTION V

DISCUSSION

SCENE TYPE

Training Scene Type had a number of important effects both on training and on transfer trials. Pitch control was better during training with the Landscape and that advantage carried over to the transfer trials. The mean advantage in transfer for training with the Landscape was greater than two degrees of RMS error, a difference that could have substantial impact on the effectiveness of an air-to-surface strike. Similar effects were found with altitude control, although some of the trends failed to achieve normally accepted levels of statistical significance.

It was also apparent that pilots who practiced the task with both the Landscape and the Grid Pattern had pitch and altitude error scores that were very similar to those of pilots who learned only with the Grid Pattern. Training scene type had some performance effects but no differential transfer effects on lineup error in the dive to the target or on lateral bomb miss distances. Transfer scene type had some performance effects that support and extend the interpretation of the performance effects due to training scene type.

PERFORMANCE EFFECT OF SCENE CONTENT. The powerful performance effect of scene content was unexpected, and its source remains obscure. Presumably, some specific visual factors rather than the sheer amount of scene detail led to this result. The strongest performance effects of scene content appeared in the dive to the target, a task segment that appears at first glance to depend only a little on outside visual reference. The crucial judgments in the dive to the target are based on instrument readings or on the position of the target in relation to the bombsight markings. However, it is possible that the impoverished nature of the schematic Grid Pattern produced some disruptive effects. It was apparent that the white lines and the white target cube overlaid on the green ground plane produced a percept that was only marginally stable. Neither the lines nor the target was perceptually fixed to the ground plane, and they occasionally gave the impression that they were floating in space. While it seems improbable that ground cues are used for crucial altitude judgments or pitch judgments, perceptual instability could bias judgments and encourage a drift from the appropriate path at any time a pilot is not monitoring the bombsight and the instruments.

The possibility that scene content would affect performance in the cone segment of the task, although thought unlikely, had been considered prior to the experiment. Visual references are required to maintain a distance of approximately two miles from the target in the circling part of the cone pattern and to judge offset angles from the run-in line for initiation of roll-in and roll-over. It would be consistent with the theoretical development of Gibson (1950) and the empirical work of Hammerton and Tickner (1968) to suggest that the Landscape provided better references for scaling from realistic size and position relationships of houses, towns, roads, and fields. Nevertheless there was no evidence in the cone performance data that scene content affected that segment of the task.

The experiment was not designed to permit comparative analysis of the effects of specific visual features. At the outset it seemed sufficient to test the gross effects of scene content, particularly in view of the expectation that it would have little effect. Our result contrasted strongly with our expectation, and one or more of the many differences between the two scenes could have contributed to the effect. However, it is not possible to ascertain from the data whether the Landscape represents an optimum design or whether further enhancement (or even some degradation) would be desirable. The design of visual scenes for simulators presents almost limitless possibilities for variation. There is at present no empirical and little theoretical work to guide the design of visual displays for training simulators. The substantial performance differences found in both the training and transfer phases indicate the value of further research aimed at identifying the specific visual features and processes that can influence performance.

Two possible explanations for the scene-content effects have emerged from the prior discussion. The first is based on notions of scaling (Gibson, 1950) and suggests a need for topographic features to be distributed throughout the visual scene. Representation of familiar features of known size may have assisted altitude or distance judgments, or may have enabled judgment of motion or relative motion through enhanced motion parallax, streaming, or looming. The second explanation is based on a notion of perceptual instability and suggests a need for more detail in and near the target. Shading, shadows, texture, or the location of topographical features in the target area may stabilize the percept. An experimental contrast of these notions may provide a useful step toward defining scene content requirements.

During transfer, pitch control in the dive was good with the new River Valley scene but poor with the Grid Pattern. Altitude control in the dive showed a different pattern with the best scores coming from the Landscape with the 360 degree run-in line, and the worst from the new River Valley scene with the 360 degree run-in line. The absolute longitudinal bomb miss

distance also showed that pilots experienced some difficulty on the River Valley scene with the 360 degree run-in line. These data indicate that even after substantial training, a pilot's performance can be affected by specific features of a visual scene.

The poor altitude control in the cone segment with the River Valley scene and 180-degree run-in heading may have resulted from the need for pilots to fly over 4000-foot mountains. In contrast, the alternate run-in heading for this new scene, in which pilots flew around the end of the mountain range, resulted in poorer dive altitude control. However, pitch control was better with both run-in headings on this scene than with the other scenes. Again it is difficult to identify or even to speculate about the processes that affect performance, but it is clear that variations in scene detail have substantial and complex effects.

The only other noteworthy finding in relation to the effects of scene content on altitude error in the dive was the difficulty on the first transfer trial with the familiar Landscape scene using the new run-in line. This was always the first transfer trial, and pilots presumably had some early difficulty in adjusting to one or more new features following extended practice with a consistent set of conditions. Although pilots had been briefed on the changes from training to transfer, most had difficulty making the transition. However, the brevity of this disruption seems more surprising than its occurrence. There was no further disruption beyond the first six-trial transfer block. Brief transitional disruptions have been observed elsewhere (Hennessy, Lintern & Collyer, 1981), and suggest that it is possible to adapt quickly to new demands once basic skill patterns have been established.

In training, inexperienced pilots performed better in terms of RMS lateral error on the Grid Pattern than they did with the Landscape ($p < .10$). This is presumably due to the strong lineup cues provided by a corridor of parallel lines that led to the target along the run-in line. Although the major lines of the Landscape were parallel to or transverse to the run-in line, they did not form as a well-defined corridor as in the Grid Pattern.

Lineup cues continued to aid performance when they were present during transfer. Run-in headings of 360 degrees or 180 degrees were generally associated with good lineup performance in the dive to the target. At this stage even the inexperienced pilots were apparently able to use the more subtle lineup cues available in the familiar Landscape scene. However, all pilots experienced difficulty in maintaining a good run-in line with a heading of 225 degrees. Some difficulty in detecting the diagonal run-in line may also account for the large altitude errors found in the cone segment on trials with that run-in

heading. Lateral miss distances for bomb drop error scores were also larger. This was particularly true with the Grid Pattern following approaches along headings of 225 degrees and possibly points to some added difficulty in maintaining a path that is diagonal to compelling lineup cues. Although there was a diagonal symmetry on the 225-degree heading, pilots were apparently unable to use it effectively as a lineup cue.

DIFFERENTIAL TRANSFER EFFECTS OF SCENE CONTENT. Transfer from the Landscape was better than transfer from the Grid Pattern scene or from both scenes. The better performance on the Landscape scene during training indicates that pilots use some features of that scene not present in the Grid Pattern. It is puzzling that pilots can transfer from a scene that has seemingly important visual features, to one that lacks them, without some disruption. These results suggest that this air-to-ground bombing task is, in early learning, predominantly a closed-loop skill in which information from the visual scene is used to guide behavior. As training progresses, the task apparently evolves to become a more open-loop skill in which the motor patterns are executed with less dependence on the visual scene. It should be noted that pilots transferring from the Landscape performed better on the Grid Pattern than did pilots who were trained on the Grid Pattern. This might be considered as the most important finding of this experiment. The fact that various details of a visual scene can affect performance while they are present was interesting in itself but does not necessarily imply that they are important for training. The experimental literature is replete with examples of differential effects during training that disappear on transfer to a criterion condition (Lintern & Gopher, 1980). However, we have shown here that training on a visual scene with specific cultural features has a strong and lasting effect on transfer to a variety of visual conditions.

The superiority in transfer to the Grid Pattern of those trained with the Landscape is intriguing. The similarity principle of transfer (Holding, 1976) would predict that transfer between markedly different scenes would be poorer (or certainly no better) than transfer between identical scenes. Poulton (1974) has discussed several instances of asymmetrical transfer between control-system variations that involve departures from the law of similarity. Our scene-detail transfer data indicated that this asymmetrical pattern can also occur as a result of variations in visual scene content, and further suggest a progression towards independence from external visual information.

There is, however, some evidence that performance is not entirely independent of information from the external visual scene even after extended training. As previously noted, the best dive pitch control in transfer was associated with the River Valley scene. Presumably, some visual cues in this scene,

that were not available in the others assisted the pilots. Bombing accuracy might be similarly influenced by topographical features in real flight, so that accuracy might suffer in attacks on relatively featureless environments. However, a significant interaction of training scene type and transfer scene type indicated that there was no transfer advantage for the River Valley after training on the Landscape scene. These observations indicate a specific potential advantage for simulator training. We tentatively conclude that training on a carefully designed visual scene (possible only in a simulated environment) can make pilots resistant to performance decrements that might normally accompany operations in environments that have inadequate visual cues.

There was no differential advantage during transfer as a result of training with stronger lineup cues.

The effects of scene content on transfer were unexpected. Data from landing studies were cited in the introduction of this report to support our expectation that scene content would have little impact on transfer. Nevertheless, the behavioral research program at the VTRS is oriented towards reexamining these types of issues in with different flight scenarios. The difference between the landing and the air-to-ground attack results verifies the need for this approach.

PERFORMANCE EFFECTS OF SCENE VARIETY. Prior research led us to predict that training performances with two scenes would be poorer than with one. Some limited support was obtained from the lateral bomb miss distance which showed poorer training performances for pilots trained with both visual scenes. In addition this group showed very poor lineup performance during the dive to the target on their first six training trials.

The effects on altitude and pitch control in the dive to the target were more consistent but were also more difficult to interpret. The performances of pilots trained with both the Landscape and Grid Pattern scenes were similar to those of pilots trained with only the Grid Pattern. Explanations based on content or on variety might be suggested, but this pattern of results does not seem to conform to either. In fact, the facilitating effect of training with the Landscape scene to transfer performance with the Grid Pattern (noted previously) should have been present during training for the pilots trained with both scenes, but did not occur.

Under a scene content explanation we would expect those trained with both scenes to do well when flying with the Landscape scene and poorly when flying with the Grid Pattern. Under a scene variety explanation we would expect some performance decrement with each scene in relation to performances by pilots who were trained with only one or the other. Our results did not support either explanation;

performance with the Landscape scene was poor under varied training, and was surprisingly similar to performances with the Grid Pattern under either varied or consistent training. Thus, alternation between the Landscape and Grid Pattern scenes disrupted the potential of the Landscape scene to aid performance during training.

Interference between psychomotor tasks has been studied extensively by Lewis and his associates (Lewis & Shephard, 1950; Lewis, McAllister & Bechtoldt, 1953a; Lewis, McAllister & Bechtoldt, 1953b; Lewis & Miles, 1956). It is clear from this work that performance of one task can disrupt later performance on another. However, that research was based on reversals or incompatible changes in stimulus-response relationships. Our experiment allowed identical responses to different stimuli, a situation that should allow easy transition between tasks (Holding, 1976). Interference as a result of changes only in visual information appears to have no precedent in psychomotor research. Our data do suggest that scene variation in early learning can interfere with the acquisition of bombing skills.

DIFFERENTIAL TRANSFER EFFECTS OF SCENE VARIETY. There were no noticeable differential transfer effects of scene variety during training on lateral bomb miss distance or on lineup error in the dive to the target. As far as could be ascertained the effects of scene variety on altitude and on pitch control in the dive followed the pattern found in the training data. Thus, the procedure of switching between scenes apparently disrupted the transfer as well as the performance benefit that could otherwise accrue from the Landscape scene.

It is surprising that those trained with two scenes did not outperform the remaining pilots in the transfer phase. In particular, there was no advantage of training on two scenes for transfer to the new River Valley scene. Additionally, there was generally no disruption of the high level of performance established by pilots trained only on the Landscape scene when they transferred to a variety of scenes and run-in headings. The only disruptive effect on lineup during the dive to the target with the first block of transfer trials (in which a new run-in heading was used), and this effect was apparent for all groups of pilots.

These trends indicate that variation of scenes or of localities offers no particular advantage in early training and may in fact slow learning. Thus, early training should be consistent. However, the initial transfer difficulty with the new run-in heading suggests that brief exposure to new run-in headings should be planned near the end of training. Initial and transient decrements during transfer as noted by Hennessy, et al. (1981), further attest to the need for at least a few training trials with the full range of task variations that will be encountered after training. This strategy would seem to

promote the specific advantages of both consistent and varied training, those being, respectively, faster initial learning and later development of general schema that can be applied to a wider range of task variations.

The possibility that task variations can impede early learning points to a distinct advantage inherent in training with simulators versus training with aircraft. Variations and constancies can be reliably scheduled in simulator training, whereas they are left to chance and logistics in aircraft training. A further potential advantage is that representation of an operational target area may permit a pilot to practice specific attacks prior to his first flight into that area. This could be of enormous assistance in wartime operations where a first pass can be crucial.

The failure of scene variation during training to have the expected impact on transfer is puzzling, especially in view of substantial research that has shown a facilitating effect. The nature of the task may be significant, and it is apparent that our air-to-ground task is very different to the laboratory tasks that have been used to test the effects of variety on transfer. Nevertheless, variety is an important issue in simulation training and variations, other than of scene content, could be considered. The ability to handle a range of weapon-delivery tasks and a range of environmental conditions is crucial to an attack pilot. Simulation affords an opportunity to expose trainees to a wide variety of situations. Further research is needed to establish the useful dimensions and limits of task variation for training purposes.

It may also be premature to conclude that variation between scenes will seriously disrupt training. The two training scenes were as different as is feasible in the VTRS, and the Grid Pattern was unnatural and impoverished. Variation between realistic but otherwise different scenes may not produce the same result. Thus, further research in the effects of scene variation is also warranted.

AUGMENTED FEEDBACK

Augmented Feedback had strong effects on aircraft control in the dive towards the target. The benefit to altitude and pitch control as a result of training with Augmented Feedback impacted training, probe and transfer trials. The advantage was substantial and would be expected to influence miss distances by more than 100 feet. This value is meaningful and could represent a difference between considerable and minor damage to a target.

PERFORMANCE EFFECTS OF AUGMENTED FEEDBACK. The effect of augmented feedback on altitude and pitch control during training partially validates the use of this type of augmenting cue.

With very little prior research to use as a guide, the augmented feedback had been implemented with only limited pretesting to confirm that it would be effective. Other augmented feedback research that has failed to show any performance effect during training has also failed to show any differential transfer effect (Archer, Kent & Mote, 1956; Karlin, 1960; Micheli, 1966). In simulation research, Hughes et al. (1979), tested a bomb prediction cue for air-to-surface weapon delivery training and found that the cue did not help either training or transfer performances. It seems inherent in the concept of augmentation that the supplementary cues aid performance when they are present. Supplementary cues that do not aid performance during training usually will not provide a valid test of augmented feedback as an instructional variable.

There were no significant effects on lateral error or on lineup performance. However, there were scene content effects which suggest that lineup could be improved with appropriate augmenting cues. The postulated corridor effect in the low-detail scene for the 360 degree run-in heading suggests that an artificial visual corridor could be a useful augmenting cue for realistic scenes in which lineup remains a problem. Adaptive withdrawal of the artificial corridor might permit a performance advantage to transfer.

DIFFERENTIAL TRANSFER EFFECTS OF AUGMENTED FEEDBACK. Performance advantages due to augmented feedback carried over from training trials to transfer trials. Transfer performances of pilots at both levels of experience were aided on dive altitude control, but only the inexperienced pilots benefited on dive pitch control. In contrast, more experienced pilots performed better on longitudinal bomb miss distance as a result of augmented-feedback training while inexperienced pilots performed more poorly.

The varying patterns of results may be due to differing sensitivities of performance measures to different stages of learning. Although this phenomenon has not been well studied in the psychomotor domain, it has been observed during acquisition of other tasks (Lesgold, 1983). In the air-to-ground bombing task examined here, pilots presumably assert control in a progressive fashion over various task dimensions. Thus, experienced pilots might not profit from augmented-feedback training on dive pitch control if they are already accomplished on that dimension. However, it appears more difficult to establish good dive altitude control, so that pilots at both levels of experience probably needed to improve on this dimension, and could thereby profit from augmented-feedback training.

We speculate that the relative disadvantage in longitudinal bombing accuracy shown by the inexperienced pilots following augmented-feedback training was due to their concentration on

dive parameters at the expense of release parameters. The supplementary cues had been designed to draw attention to dive parameters. More experienced pilots could be expected to distribute their attention more effectively, and therefore avoid a decrement in bombing accuracy.

Nevertheless, we do not regard the reaction of the inexperienced pilots as a problem. The augmenting cues were designed under the assumption that early control should be established over dive parameters and that terminal performances should be deemphasized. A more common example of this learning strategy occurs with typing where beginning students concentrate on developing crucial sub skills and ignore terminal performance until they have done so. Those who learn to type by concentrating at the outset on terminal performance are unlikely to develop a high level of skill. For our bombing task, we had assumed that the augmented feedback would modify the sequence in which aspects of the task were learned, and that the modified sequence would be more efficient. Trends in data from more experienced pilots suggest that inexperienced pilots who trained with augmented feedback could resolve their problems with bombing accuracy, and in the process would develop more effective habits that would produce better and more reliable bombing scores.

Inspection of the probe dive-pitch control data further illuminates the effects of augmented feedback on skill development. Learning patterns were modified by augmented-feedback training and by scene-type in training. Improvement was evident for the augmented-feedback groups, but not for the nonaugmented-feedback groups. Even the more experienced pilots appeared to benefit from augmented feedback training, although the effect was apparently short-lived because the subsequent transfer data did not show a similar advantage.

The failure of the nonaugmented-feedback groups to improve their dive pitch control was puzzling. Learning on this dimension is apparently possible in an aircraft without augmented feedback because the more experienced pilots performed better on this measure. Further examination of scene content effects in the probe trials indicated that nonaugmented-feedback groups did not improve their pitch control if they trained with the Grid Pattern or with both scenes. Thus, augmented feedback seems essential for training the task with impoverished scenes, and may overcome some of the disadvantage of using such scenes.

One potential problem with augmented feedback is that the advantage may not generalize well to new run-in headings. Following augmented-feedback training, altitude control was better during transfer to task variations that included the training run-in heading but not during transfer to task variations with different run-in headings. This should not be considered as a serious objection at this stage. It is likely

that augmented-feedback training schedules could include different headings so that the benefits would be more easily generalized.

Our data have shown that a simple set of augmenting cues can have powerful and complex effects on several task dimensions in training, probe, and transfer. Nevertheless, augmented feedback should be used with care. Poorly designed augmented feedback could teach inefficient strategies or habits as effectively as well designed augmented feedback can teach efficient ones. This experiment has shown that augmented feedback can be a powerful instructional tool. However, it cannot be applied systematically until we know more about the task dimensions that should be targeted, the types of supplementary cues that should be used, and the types of adaptive withdrawal schedules that should be employed.

PRIOR BOMBING EXPERIENCE

There were surprisingly few significant effects of pilot experience, although several trends were nearly significant and in the expected direction. The primary reason to include pilot experience in the experiment was to examine its interactions with other factors. There were several such interactions, and they have been noted and discussed in previous sections. To summarize these effects, it appears that inexperienced pilots benefit more from the assistance of high scene content and augmented feedback. Nevertheless, the moderately experienced pilots also benefit from these treatments, albeit in different ways or to a lesser extent. There is no systematic evidence in these data that pilots with no air-to-ground attack experience should be treated differently from those who have a moderate amount.

ISSUES FOR EXPERIMENTAL DESIGN

Statistical power remained distressingly low with some dependent measures despite the care directed toward providing satisfactory power for the statistical tests. The attempt to provide a covariate was unsuccessful, and the strategy of relaxing the significance criterion to 0.10 improved power only marginally. We must assume that some real effects failed to achieve appropriate levels of statistical significance and were therefore not discussed. This is an unsatisfactory situation, but the most direct method of increasing power, that being to increase the number of experimental subjects, would have increased the cost of an already expensive experiment beyond reach.

Our attempt to provide a satisfactory covariate failed, and further illustrates the danger of generalizing from one flight task to another. Unfortunately, there were few data and only meager theory to guide our selection, although recent data from

Kruk, Regan, Beverley and Longridge (1983) identifies some visual tracking tests that could be effective. The challenge of providing satisfactory power for transfer-of-training experiments must be faced and covariates and more reliable performance measures offer the greatest potential for doing so without substantially increasing the costs associated with the experiment. Further exploration of covariates and performance measures is recommended.

Probe methodology was proposed by Hughes, et al. (1982), as a means of monitoring transfer performance throughout training. In that experiment, and another by Smith, Pence, Queen, and Wulfeck (1974), there was no other test of transfer. In general, our probe and transfer results are consistent with each other. However, examination of the probe data for the pitch control measure did lead to some useful observations on the effects of augmented feedback. Thus, there appears to be some value in using probe and transfer tests in the same experiment.

Nevertheless, there is a danger that probe trials could impact the course of learning. For example, Hagman (1983) has shown that the scheduling of feedback and no-feedback trials can affect the short- and long-term retention of a discrete motor skill. Test or probe trials could enhance the training effects of augmented feedback by discouraging dependency on the supplementary cues much in the way that adaptive withdrawal appears to. Alternatively, frequent transition from training to probe trials might have some disruptive effect on probe performance so that it would not accurately reflect the course of learning. Thus, probe methodology should be used with caution. It may be useful in exploratory experiments, where it could uncover trends not evident in transfer data, but it should be avoided in research that is intended to determine procedures for an operational training program.

SECTION VI

CONCLUSIONS

Scene content had an unexpected but strong and consistent effect on performance and on transfer. The Landscape scene (a relatively highly detailed scene) was generally superior to the Grid Pattern scene although the Grid Pattern proved to be superior on some measures. While it was not generally possible to determine with confidence scene features that contributed to this effect, some likely candidates were identified.

The most intriguing observation on scene content was that pilots who learned the task with the Landscape could later perform well with the Grid Pattern. However, those trained on [illegible] Grid Pattern never exhibited the high level of performance [illegible] by those trained on the Landscape. Thus, some of the [illegible]res of the Landscape seem essential for early learning but not for later performance. Their value for later learning could not be determined.

The scene-content issue is one of the most crucial for modern training simulators. Our data are the first to show that scene content affects learning of flight skills. Further research to identify specific visual features that do impact learning is essential. In the meantime, simulator training of air-to-ground attack should be conducted with visual scenes that at least have features similar to those in our Landscape. Whether all of those features are necessary or sufficient for efficient training should be determined in future research.

Variety was raised as a training issue specifically because modern simulators can provide enormous variety at little additional cost. Scene variety in training did not generally benefit transfer, and there is a distinct possibility that it can interfere with early learning. However, transient disruptions in performance at transfer suggested that brief experiences with a wider range of scenes towards the end of a constant training regimen could be useful.

Thus, we recommend that air-to-ground attack be taught initially in a simulator with only one scene and run-in heading but that a variety of scenes and run-in headings be introduced just prior to transfer. A similar approach might be used in the aircraft if the option exists. In addition, it may be possible to simulate attacks on actual targets. This might increase the effectiveness of a pilot's first pass at that target. The effect on training of varying other factors of the task could

also be examined. Variations in environmental conditions and ordnance delivery modes are good candidates for further research.

Augmented feedback proved to be a potent instructional variable but one that showed complex effects. It helped inexperienced pilots with their dive pitch control and helped both groups of pilots with their dive altitude control. The data further indicated that augmented feedback helped more experienced pilots with their longitudinal bomb miss distance. Thus, the effects of augmented feedback are pervasive and progressive. It would appear to be useful at least for primary and intermediate instruction.

As yet little is known about augmented feedback. The bulk of the previous research has involved simple tracking tasks. Our data have extended the knowledge about how it might be applied to air-to-ground attack training and what benefits can be anticipated. Nevertheless, further research is necessary to determine optimum designs for the augmenting cues and appropriate adaptive withdrawal schedules. In addition, we need to explore the range of its effects with other flight tasks.

There were several interactions of pilot experience with the experimental variables. In general, the inexperienced pilots suffered most from low scene detail and gained most from augmented feedback. Nevertheless, the moderately experienced pilots were also affected by these variables. Thus, there is no generalizable evidence that pilots with no experience in air-to-ground attack should be treated differently during training from pilots who have some experience in air-to-ground attack.

The experimental design issue of statistical power was also considered. Statistical power remains a problem area. Despite our best efforts, power remained low on some important performance measures. The most obvious means of increasing power, that being to increase the number of experimental subjects, will often be considered too expensive. The options of finding satisfactory covariates and of improving performance measures should be pursued.

REFERENCES

Adams, J.A. Multiple versus single problem training in human problem solving. Journal of Experimental Psychology, 1954, 48, 15-18.

Archer, E.J., Kent, G.W., and Mote, F.A. Effect of long-term practice and time-on-target information feedback on a complex training task. Journal of Experimental Psychology, 1956, 51, 103-112.

Berry, F. and Lintern, G. Briefing: Ordnance delivery in the visual technology research simulator, thirty-degree bomb. Orlando, FL: Naval Training Equipment Center, Visual Technology Research Simulator, Unpublished document, 1982.

Ciavarelli, H.P., Williams, A.M., and Brictson, C.A. Development and application of air combat performance assessment methods. Proceedings of the First Symposium on Aviation Psychology. The Ohio State University, Department of Aviation, Columbus, OH: 1981.

Cohen, J. Statistical power analysis for the behavioral sciences. New York, NY: Academic Press, 1977.

Crafts, J. Routine and varying practice as preparation for adjustment to a new situation. Archives of Psychology, 14 (91), 1927.

Dashiell, J.F. An experimental isolation of higher level habits. Journal of Experimental Psychology, 1924, 7, 391-397.

Duncan, C.P. Transfer after training with single versus multiple tasks. Journal of Experimental Psychology, 1958, 55, 63-72.

Ellis, H.C. Transfer and retention. In M.H. Marx (Ed.), Learning: Processes. London, England: Collier-MacMillan, 1969.

Gibson, E.J. Principles of perceptual learning and development. New York, NY: Appleton-Century-Crofts, 1969.

Gray, T.H. and Fuller, R.R. Effects of simulator training and platform motion on air-to-surface weapons delivery training. Brooks Air Force Base, TX: USAF Human Resources Laboratory, AFHRL-TR-7729, 1977.

Hagman, J.D. Presentation- and test-trial effects on acquisition and retention of distance and location. Journal of Experimental Psychology, 1983, 9, 334-345.

Hammerton, M. and Tickner, A.H. An investigation into the effects of stress upon skilled performance. Ergonomics, 1968, 12, 851-855.

Hennessy, R., Lintern, G., and Collyer, S. Unconventional visual displays for flight training. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 78-C-0060-5, 1981.

Holding, D.H. An approximate transfer surface. Journal of Motor Behavior, 1976, 8, 1-9.

Hughes, R.G., Lintern, G., and Wightman, D.C. Applications of simulator freeze to carrier glideslope tracking instruction. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 78-C-0060-9/AFHRL-TR-82-3, 1982.

Hughes, R.G., Paulsen, J., Brooks, R., and Jones, W. Visual cue manipulation in a simulated air-to-surface weapon delivery task. Proceedings of 1st Interservice/Industrial Training Equipment Conference, USAF Human Resources Laboratory, Flying Training Division, NAVTRAEQUIPCEN IH-316, 1979.

Jones, M.B. Behavioral Science 511 (Statistics and Research Design). Hershey, PA: Department of Behavioral Sciences, Milton S. Hershey Medical Center, Pennsylvania State University, Unpublished.

Karlin, L. Psychological study of motor skills: Phase 1. Pt. Washington, NY: Naval Training Device Center, Technical Report NAVTRADEVCEN 558-L, 1960.

Kelley, R.S., Poulter, D.C., and Castore, C.H. Simulated A-10 combat environment. Proceedings of the Image Generation/Display Conference II, Williams AFB, AZ: USAF Human Resources Laboratory, Flying Training Division, 1981.

Kruk, R., Regan, D., Beverley, K.I., and Longridge, T. Flying performance on the advanced simulator for pilot training and laboratory tests of vision. Human Factors, 1983, 25, 457-466.

Lesgold, A.M. Acquiring expertise. Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh. UPITT/LRDC/ONR/PDS-5, 1983.

Lewis, D., McAllister, D.E., and Bechtoldt, H.P. Correlational analysis of the learning and relearning of four different tasks on the modified Mashburn apparatus. _Journal of Psychology_, 1953a, _36_, 83-109.

Lewis, D., McAllister, D.E., and Bechtoldt, H.P. Correlational study of performance during successive phases of practice on the standard and revised tasks on the SAM complex coordinator. _Journal of Psychology_, 1953b, _36_, 111-126.

Lewis, D. and Miles, G. Retroactive interference in performance on the star discrimeter as a function of amount of interpolated learning. _Perceptual and Motor Skills_, 1956, _6_, 295-298.

Lewis, D. and Shephard, A.H. Devices for studying associative interference in psychomotor performance. I. The modified Mashburn apparatus. _Journal of Psychology_, 1950, _24_, 35-46.

Lintern, G. Transfer of landing skill after training with supplementary visual cues. Champaign, IL: University of Illinois, Department of Psycnology, TR Engineering Psychology-78-3/AFOSR-78-2, 1978.

Lintern, G. Transfer of landing skill after training with supplementary visual cues. _Human Factors_, 1980, _22_, 81-88.

Lintern, G. and Gopher, D. Adaptive training of perceptual-motor skills. Issues, results and future directions. _International Journal of Man-Machine Studies_, 1978, _10_, 521-51.

Lintern, G. and Kennedy, R.S. A video game as a covariate for carrier landing research. Proceedings of the Eighth Psychology in the Department of Defense Symposium. Colorado Springs, CO: USAF Academy, 1982.

Lintern, G.T. and Roscoe, S.N. Visual cue augmentation in contact flight simulation. In S.N. Roscoe, _Aviation Psychology_. Ames, IA: The Iowa State University Press, 1980.

Micheli, G.S. Augmenting feedback and transfer of training. Orlando, FL: Naval Training Device Center, Technical Report NAVTRADEVCEN 1H-41, 1966.

Morisett, I.J. and Hovland, C.I. A comparison of three varieties of training in human problem solving. _Journal of Experimental Psychology_, 1959, _58_, 52-55.

Naval Air Training Command, Flight Training Instructions TA-45, Parts VII and VIII. Corpus Christi, TX: Naval Air Station, 1980.

Poulton, E.C. Tracking skill and manual control. New York, NY: Academic Press, 1974.

Roscoe, S.N. and Childs, J.M. Reliable, objective flight checks. In S. N. Roscoe. Aviation Psychology. Ames, IA: The Iowa State University Press, 1980.

Schmidt, R.A. A schema theory of discrete motor skill learning. Psychological Review, 1975, 82, 225-260.

Schneider, W. and Fisk, A.D. Attention theory and mechanisms for skilled performance. Champaign, IL: Psychology Department, University of Illinois HARL-ONR-8201, 1982.

Shapiro, D.C. and Schmidt, R.A. The schema theory: Recent evidence and developmental implications. In J.A.S. Kelso and J.E. Clark (Eds.). The development of movement control and coordination. New York, NY: John Wiley and Sons, 1982.

Shiffrin, R.M. and Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 1977, 84, 127-790.

Smith, R.L., Pence, G.G., Queen, J.E., and Wulfeck, J.W. Effect of a predictor instrument on learning to land a simulated jet trainer. Inglewood, CA: Dunlap and Associates, 1974.

Vreuls, D. and Sullivan, D.J. Ground attack tasks and performance measurement for future Visual Technology Research Simulator studies. Thousand Oaks, CA: Vreuls Research Co., Report No. N61339-80-C-0131, 1981.

Waag, W. L. Training effectiveness of visual and motion simulation. Williams AFB, AZ: USAF Human Resources Laboratory, Operations Training Division, AFHRL-TR-79-72, 1981.

Westra, D.P. Simulator design features for carrier landing: II. In-simulator transfer of training. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 81-C-0105-1, 1982.

Westra, D.P., Simon, C.W., Collyer, S.C., and Chambers, W.S. Simulator design features for carrier landings: I. Performance experiments. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 78-C-0060-7, 1981.

Wightman, D. Part-task training stragegies in simulated carrier landing final approach training. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN IH-347, 1983.

APPENDIX A

STATISTICAL POWER

Power calculations for two previous studies of air-to-ground attack are shown in Table A-1. In both studies bomb miss distance was used as a performance measure. Discussions with pilots in preparation for our experiment had indicated that bomb miss distances of 40, 70 and 100 feet represented small, medium, and large effects. Power values calculated from Gray and Fuller (1977) using a significance level of 0.05 were 0.18, 0.47 and 0.78 for small, medium, and large effects. Similar calculations on data from Hughes et al. (1979) gave power values of 0.10 and 0.23 and 0.42. These values seem undesirably low.

TABLE A-1. POWER CALCULATIONS ON BOMB MISS DISTANCE DATA FROM SIMULATION STUDIES OF AIR-TO-SURFACE WEAPONS DELIVERY RESEARCH

	Small Effect	Medium Effect	Large Effect
	(40 ft.)	(70 ft.)	(100 ft.)
Calculations from Gray and Fuller (1977)			
MSE = 4715 ft.			
Power (Total N = 16, p = .05)	0.18	0.47	0.78
Total N required for Power of 0.95 (p=.05)	142.00	48.00	24.00
Power for N = 32 (p = .05)	0.34	0.81	0.99
Power for N = 32 (p = .10)	0.48	0.90	0.99+
Calculations from Hughes et al., (1979)			
Standard Deviation = 104 ft.			
Power (Total N = 16, p = .05)	0.10	0.23	0.42
Total N required for Power of 0.95 (p = .05)	324.00	108.00	54.00
Power for N = 32 (p = .05)	0.17	0.44	0.76
Power for N = 32 (p = .10)	0.27	0.59	0.87

Selection of an appropriate level of power is a matter of judgment, as is selection of a significance criterion. However, conventions have been established for significance criteria, whereas they have not been established for power. Cohen (1977) offers the value of 0.80 as a sensible convention and thereby implies that the seriousness of a type I error (falsely rejecting the null hypothesis) is four times that of a type II error (falsely accepting the null hypothesis) when this power criterion is coupled with a significance criterion of 0.05.

In applied research the 1:4 seriousness ratio for the two types of errors may not be appropriate. A type II error in this experiment could lead to a decision to eliminate some effective simulator features, thereby resulting in a less effective device. A type I error, on the other hand, could result in an unnecessarily expensive, albeit effective device. Initially it seemed desirable to establish a power criterion of 0.95 with a significance criterion of 0.05. This procedure equates the seriousness of type I and type II errors.

The Gray and Fuller (1977) data indicate that 142 subjects would be required for the most desirable situation, that being a power of 0.95 for a small effect with a significance level of 0.05. The high per subject cost for this experiment forced a more detailed consideration of the number of subjects that should be used. Further analysis of the data indicated that 32 subjects would result in a power of 0.95 for a medium effect with a significance criterion of 0.10. This number of subjects permitted an experimental design that could test the four factors of interest with two subjects in each cell. It is questionable whether the identification of a small effect would have any practical or operational value. Thus, 32 subjects were used in the experiment. This gave a sufficiently powerful test of experimental effects and also permitted a balanced experimental design.

APPENDIX B

FLIGHT PATH DEFINITION FOR BOMB DELIVERY FROM A 30-DEGREE PATTERN

Basic parameters for the task were selected to conform as closely as possible to flight training instructions for the TA-4J (Naval Air Training Command, 1980). Compromise in the selection of values for some parameters (e.g., airspeeds) was necessary to be consistent with the performance envelope of the T-2.

Details of the flight path are shown in Figures B-1 and B-2. The radius of the flight path from the abeam position to roll-in was 12,938 feet. Airspeed during this portion of the flight was to be 250 knots, and pilots were advised that a bank of approximately 30 degrees would produce an appropriate rate of turn. Altitude was to be maintained at 8000 feet.

The roll-in point subtended an angle of 30 degrees at the target with the run-in line. Pilots were advised to roll into a bank of 45 degrees and to maintain 8000 feet of altitude. The radius R of the arc defining this segment of the flight path was 5787 feet and was calculated from

$R = V/W$, where

V = velocity in feet per second, and

W = turn rate in radians per second.

The turn rate of 250 knots and a bank of 45 degrees was established empirically.

The roll-over point subtended an angle of 10 degrees at the target with the run-in line. At the roll-over point, pilots were expected to reduce power to 84%, to roll to 120 degrees of bank, to allow the nose of the simulator to drop to a 30 degree dive, and to pull the 120-mil ring of the bomb sight towards the target. There was no apparent way to define the coordinates of this segment of the flight path.

Pilots were instructed to roll the wings level on the run-in line, and in a 30 degree dive when the 120-mil ring of the bombsight lay on the target. The dive path to the target is shown in Figure B-2. It is shown as extending from 8000 feet altitude to ground level, although pilots would normally intercept it at approximately 6500 of altitude, and would pull away from it at approximately 2500 of altitude. As is apparent from Figure A-2, the intersection of the dive path with the

ground is beyond the target. That distance (1511 feet) was determined by first calculating how far the bomb would travel if released at the optimum point and with optimum velocities. The remaining values in Figure B-2 could be calculated by simple trigonometry from known values.

The flight path for the pull-up segment could not be determined precisely. A maximum loading of 4g and a minimum altitude of 1200 feet were established.

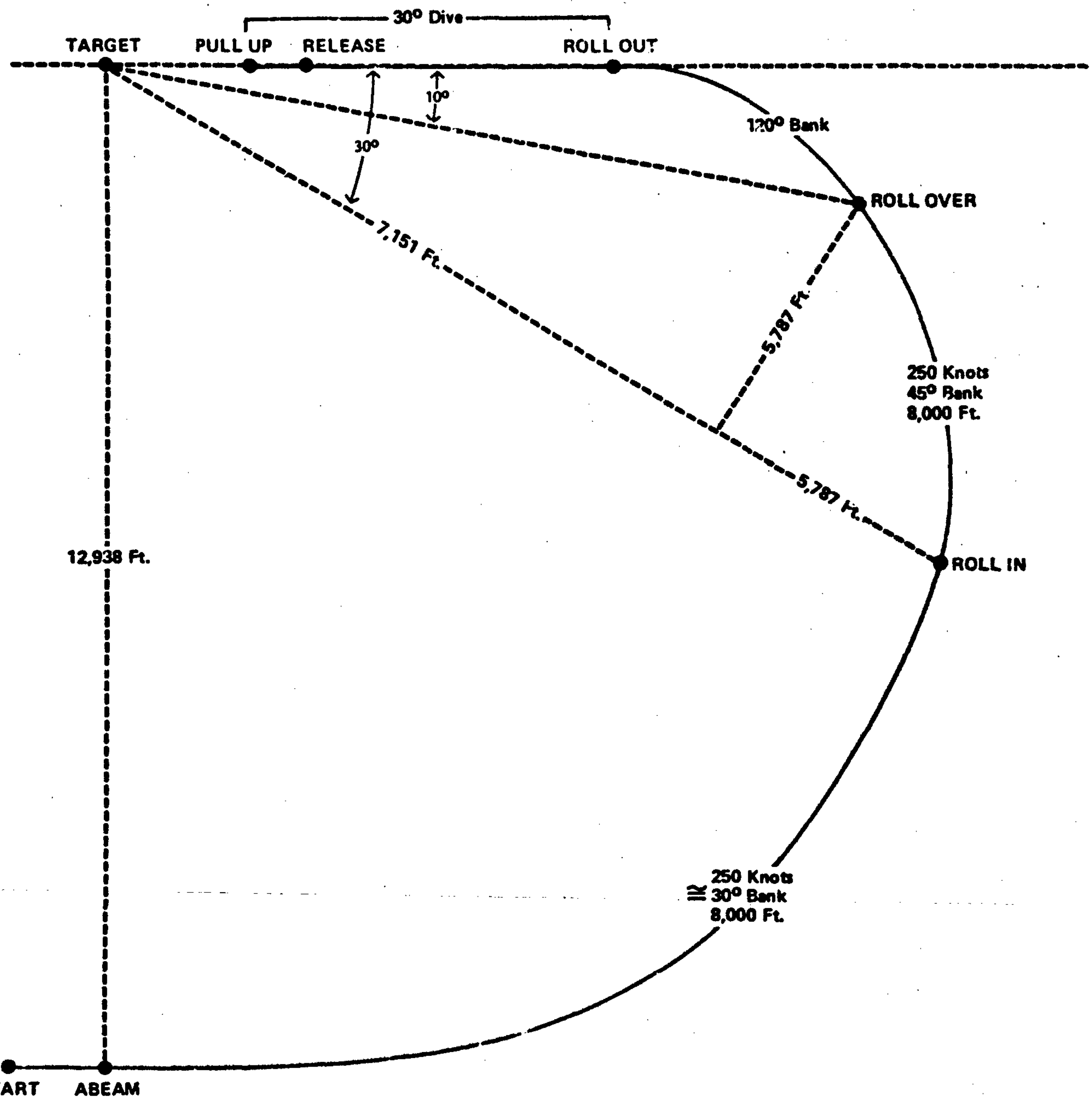


Figure B-1. Ground Track of Flight Path for the Experimental Task

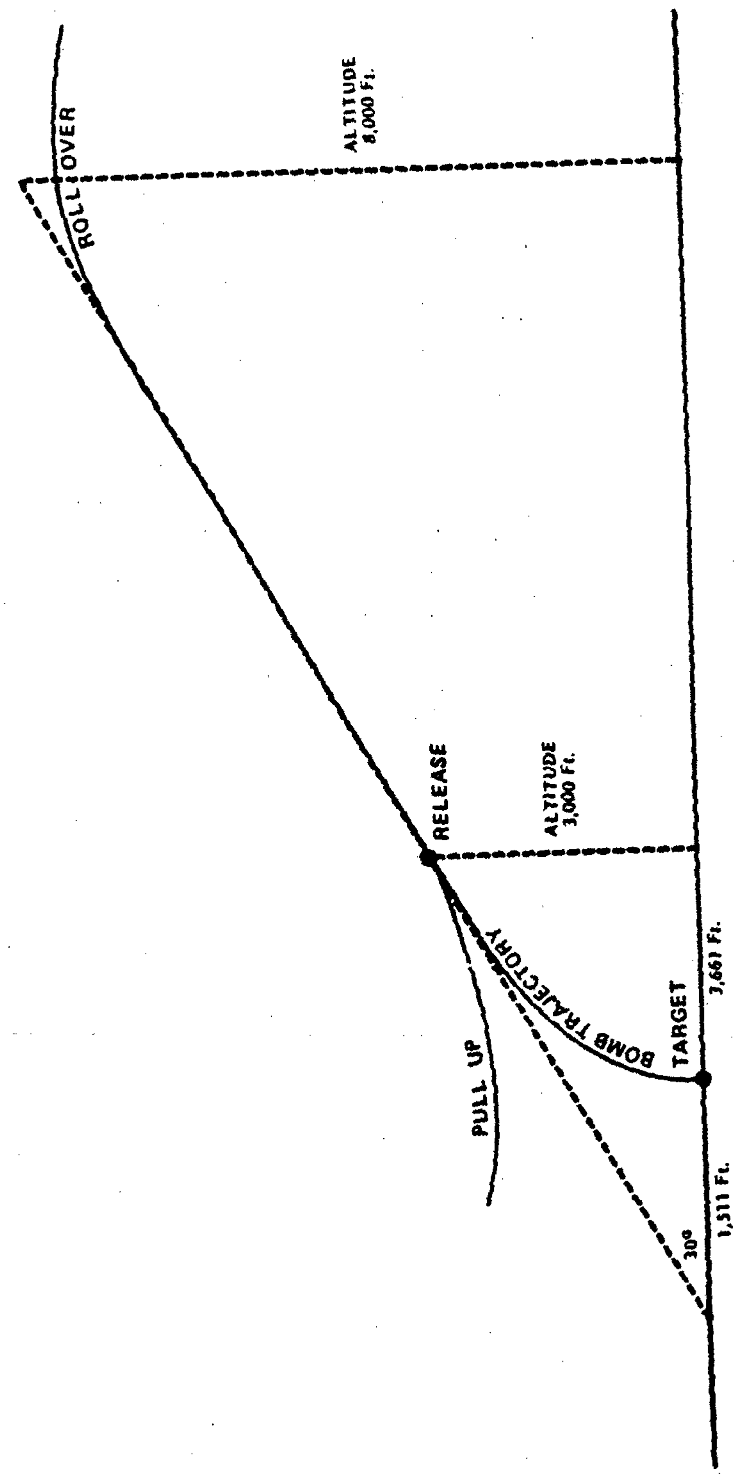


Figure B-2. Dive Path to the Target

DISTRIBUTION LIST

Naval Training Equipment Center 1
Orlando, FL 32813

Commander 3
HQ, TRADOC
Attn: ATTNG-PA
Ft. Monroe, VA 23651

Center Library 3
Naval Personnel Research and Development Center
San Diego, CA 92152

Dr. Ralph R. Canter 1
U. S. Army Research Institute Field Unit
P. O. Box 16117
Fort Harrison, IN 46216

Defense Technical Information Center 12
Cameron Station
Alexandria, VA 22314

PERI-OU 1
U. S. Army Research Inst for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

OASD (MRA - L)/Training 2
Room 3B922, Pentagon
Washington, D C 20301

Dr. Geoffrey Grossman 1
Head, Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

The Van Evera Library 1
Human Resources
300 North Washington St
Alexandria, VA 22314

Library 1
Division of Public Documents
Government Printing Office
Washington, D. C. 20402

Director 2
Training Analysis and Evaluation Group
Department of the Navy
Orlando, FL 32813

HumRRO/Western Division/ Carmel Ofc 1
27857 Berwick Drive
Carmel, CA 93923

U. S. Coast Guard HQ 1
(G-P-1/2/42)
400 Seventh St., SW
Washington, D C 20593

Personnel & Training Research Program 3
Office of Naval Research
(Code 442)
Psychological Sciences Div.
800 N. Quincy Street
Arlington, VA 22217

National Aviation Facilities Experimental Center Library 1
Atlantic City, N J 08405

AFHRL/TSZ 2
Brooks AFB, TX 78235

American Psychological Association 1
Psyc. INFO Document Control Unit Bolling AFB
1200 Seventeenth Street, NW
Washington, D C 20036

AFOSR/NL (Dr. A. R. Fregley) 1
Bolling AFB
Washington, D C 20332

Navy Domestic Tech. Transfer 1
Fact Sheet
Code E411
Naval Surface Weapons Center
Dahlgren, VA 22448

Human Factors Society 2
Attn: Bulletin Editor
P. O. Box 1369
Santa Monica, CA 90406

Center for Naval Analyses 1
Attn: Dr. R. F. Lockman
2000 N. Beauregard Sc
Alexandria, VA 22311

National Defense University 1
Research Directorate
Ft. McNair, D C 20319

Dr. J. Huddleston 1
Head of Personnel Psychology
Army Personnel Research Establishment
c/o RAE, Farnborough
Hants, ENGLAND

Commanding Officer 1
Air Force Office of Scientific Research
Technical Library
Washington, D C 20301

OUSDR&E (R&T) (E&LS) 1
CDR Paul R. Chatelier
Room 3D129, The Pentagon
Washington, D C 20301

Dr. D. G. Pearce 1
Behavioral Science Division
Defense and Civil Institute of Environmental Medicine
P. O. Box 2000
Downsview, Ontario M3M, CANADA

Dr. Jesse Orlansky 1
Science and Technology Division
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Technical Library 1
OUSDR&E
Room 3D122
Washington, D C 20301

Chief of Naval Operations 1
OP-987H
Attn: Dr. R. G. Smith
Washington, D C 20350

Commander 2
Naval Air Systems Command
AIR 340F
Attn: CDR T. Jones
Washington, D C 20361

Scientific Technical Information Office 1
NASA
Washington, D C 20546

Chief 1
ARI Field Unit
P. O. Box 476
Ft. Rucker, AL 36362

Chief of Naval Operations 1
OP-115
Washington, D C 20350

Technical Library 1
Naval Training Equipment
Center
Orlando, FL 32813

Chief of Naval Operations 1
OP-596
Washington, D C 20350

Commander 1
Naval Air Test Center
CT 176
Patuxent River, MD 20670

Office of Deputy Chief of 1
Naval Operations
Manpower, Personnel and
Training (OP-01)
Washington, D C 20350

Assistant Secretary 1
of the Navy
Research, Engineering & Systems
Washington, D C 20350

HQ Marine Corps 1
Code APC
Attn: LTC J. W. Biermas
Washington, D C 20380

Chief of Naval Operations 1
OP-593B
Washington, D C 20350

Scientific Advisor 1
Headquarters U. S. Marine Corps
Washington, D C 20380

Dr. Martin Tolcott 1
Office of Naval Research
800 N. Quincy Street
Code 442
Arlington, VA 22217

Commander 1
Naval Air Development Center
Attn: Technical Library
Warminster, PA 18974

Naval Research Laboratory 1
Attn: Library
Washington, D C 20375

Chief of Naval Education 6
and Trng Liaison Office
AFHRL/OTLN
Williams AFB, AZ 85224

Dr. Donald W. Connolly 1
Research Psychologist
Federal Aviation Admin
FAA NAFEC ANA-230 Bldg. 3
Atlantic City, NJ 08405

Chief of Naval Material 1
MAT 0722
BCT-1, Room 503
800 N. Quincy St.
Arlington, VA 22217

Commanding Officer 1
Naval Education Training
Program & Development Center
Attn: Technical Library
Pensacola, FL 32509

Commander 1
Naval Air Systems Command
Technical Library
AIR-950D
Washington, D C 20361

Chief of Naval Education 1
and Training
Code 01A
Pensacola, FL 32509

Commander 1
Pacific Missile Test Center
Point Mugu, CA 93042

Commander 1
Naval Air Systems Command
AIR 413
Washington, D C 20361

Commanding Officer 1
Naval Aerospace Medical
Research Laboratory
Code L5
Department of Psychology
Pensacola, FL 32512

Dr. Thomas Longridge 1
AFHRL/OTR
Williams AFB, AZ 85224

Dr. Kenneth Boff 1
ARAMRL/HEA
WP AFB, Ohio 45433

CAPT James Goodson 1
Code L-32
Naval Aerospace Medical
Research Laboratory
Pensacola, FL 32512

Engineering Psychology 1
Group
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Mr. Edward M. Connelly 1
Performance Measurement Assoc.
410 Pine St., S. E. 300
Vienna, VA 22180

Mr. John O. Duffy 1
U. S. Army Test
and Eval. Command
Aberdeen Proving Ground
Maryland 21005

Dr. David C. Nagel 1
LM-239-3
NASA Ames Research Center
Moffett Field, CA 94035

Federal Aviation Admin 1
Technical Library
Bureau Research & Development
Washington, D C 20590

Commander 1
Naval Weapons Center
Human Factors Branch
Code (3194)
Attn: Mr. Ronald A. Erickson
China Lake, CA 93555

Mr. Robert Wright 1
USARI Field Unit
Ft. Rucker, AL 36362

LTCOL Jefferson Koonce 1
USAFA/DFBL
USAF Academy, CO 80840

Superintendent 1
Code 602
Human Factors Engineering
Division
Naval Air Development Center
Warminster, PA 18974

Dr. Will Bickley 1
USARI Field Unit
P. O. Box 476
Fort Rucker, AL 36362

CDR W. F. Moroney 1
Systems Engineering Test
Directorate (SY 70F)
Naval Air Test Center
Patuxent River, MD 20670

CDR, Soldier Support Center 1
ATTN: ATZI-NCR-SI (MAJ Wildrick)
Ft. Benjamin Harrison
Indiana 46216

Mr. Heinz Friedrich 1
Chief, Flight Simulation Branch
Dornier GmbH
Postfach 1426
D7990 Friedrichshafen FRG

Mr. Clarence A. Fry 1
U.S.Army Human Engineering Lab
Attn: DRXHE
Aberdeen Proving Ground
Maryland 21005

Dr. Richard F. Gabriel 1
Douglas Aircraft Co.
3855 Lakewood Blvd.
Long Beach, CA 90846

Dr. Stanley Collyer 1
NAVMAT 0722
800 N. Quincy St.
Arlington, VA 22219

Dr. Genevieve Haddad 1
AFOSR/NL
Bolling AFB, D C 20332

Dr. Edward R. Jones 1
Chief, Human Factors Engr.
McDonell Douglas Corp
St. Louis, MO 63166

Mr. Charles E. Kaul 1
7101 Galgate Drive
Springfield, VA 22153

Dr. William J. King 1
Ergonomics Associates, Inc.
P. O. Box 20987
Orlando, FL 32814

Dr. Herschel W. Leibowitz 1
Professor of Psychology
614 Bruce V. Moore Building
Pennsylvania State University
University Park, PA 16802

Chief, HEL Detachment 1
Attn: DRXHE-MI (Mr. Nichols)
U. S. Army Missile Command
Redstone Arsenal, AL 35898

Mr. William A. Breitmaier 1
Code 6022
Naval Air Development Center
Warminster, PA 18974

Dr. Clyde A. Brictson 1
Dunlap and Associates, Inc.
920 Kline St., Suite 203
La Jolla, CA 92037

Dr. W. Marvin Bunker 1
General Electric Corp
P. O. Box 2500
Daytona Beach, FL 32015

Dr. Paul W. Caro 1
Seville Research Corp
400 Plaza Bldg.
Pensacola, FL 32505

Mr. Vernon E. Carter 1
Northrop Corp., Aircraft Div.
3901 W. Broadway
Hawthorne, CA 90250

Dr. Eric Haseltine 1
Hughes Aircraft Corp
Building R-1
M/S C-320
Los Angeles, CA 90009

Dr. Robert T. Hennessy 1
Committee on Human Factors
National Research Council
2101 Constitution Ave.
Washington, D C 20418

Dr. William M. Hinton, Jr. 1
Allen Corporation
3751 Maquire Blvd., Suite 270
Orlando, FL 32803

Mr. James McGuinness 1
Person-System Integration
3012 Duke Street
Alexandria, VA 22314

Dr. Melvin D. Montemerlo 1
RTE-6
NASA Headquarters
Washington, D C 20546

Dr. Frederick A. Muckler 1
Essex Corporation
741 Lakefield Rd., Suite B
Westlake Village, CA 91361

Dr. Kent A. Stevens 1
Department of Computer and Information Science
University of Oregon
Eugene, OR 97403

Dr. Paul E. Van Hemel 1
Allen Corporation of America
401 Wythe Street
Alexandria, VA 22314

Mr. Donald Vreuls 1
Vreuls Research Corporation
68 Long Court, Suite E
Thousand Oaks, CA 91360

Mr. Lee Wooldridge 1
Vreuls Research Corporation
6568 University Boulevard
Orlando, FL 32807

Dr. Laurence Young 1
Dept. of Aeronautics
Massachusetts Inst. of Tech.
77 Massachusetts Ave.
Cambridge, MA 02139

Mr. J. Thel Hooks, Jr. 1
Dunlap and Associates
920 Kline St., Suite 203
La Jolla, CA 92037

Ms. Joyce Iavecchia 1
Naval Air Development Center
Code 6022
Warminster, PA 18974

Dr. Richard S. Jensen 1
Department of Aviation
Ohio State University
Box 3022
Columbus, OH 43210

Mr. Henry R. Jex 1
Systems Technology Inc.
13766 Hawthorne Blvd
Hawthorne, CA 90250

CDR, Naval Air Force 1
U. S. Atlantic Fleet
Attn: LCDR Robert Day
Norfolk, VA 23511

Dr. Wallace W. Prophet 1
Seville Research Corp
400 Plaza Bldg.
Pace Blvd. at Fairfield Dr
Pensacola, FL 32505

Professor Stanley N. Roscoe 1
Department of Psychology
New Mexico State University
Box 5095
Las Cruces, NM 88003

R. F. Schiffler 1
ASD/ENECH
Wright-Patterson AFB
Ohio 45433

Mr. Moses Aronson 1
3705 Wilder Lane
Orlando, FL 32804

Dr. Kevin Berbaum 1
P.O. Box 435
West Branch, IO 52358

Dr. Edward A. Stark 1
Link Division
The Singer Company
Binghamton, NY 13902

Mr. John B. Sinacori 1
P. O. Box 1043
Hollister, CA 95023

Dr. Harry L. Synder 1
VPI&SU, Dept. of Indus. Engr.
and Operations Research
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Diane Damos 1
Department of Psychology
Arizona State University
Tempe, AZ 85287